#### (12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

# (19) World Intellectual Property Organization International Bureau



# (43) International Publication Date 9 August 2001 (09.08.2001)

# PCT

# (10) International Publication Number WO 01/57188 A2

(51) International Patent Classification7:

C12N

- (21) International Application Number: PCT/US01/03800
- (22) International Filing Date: 5 February 2001 (05.02.2001)
- (25) Filing Language:

English

(26) Publication Language:

**English** 

(30) Priority Data:

09/496,914

3 February 2000 (03.02.2000)

09/560,875 27 April 2000 (27.04.2000)

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US Filed on US

09/496,914 (CIP) 3 February 2000 (03.02.2000) 09/560,875 (CIP)

Filed on

27 April 2000 (27.04.2000)

- (71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14, San

Jose, CA 95117 (US). DRMANAC, Radoje, T. [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

- (74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Coh, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Published:

without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

# NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

# 1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

#### 2. BACKGROUND

5

10

15

20

25

30

35

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

# 3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-1350. The polypeptides sequences are designated SEQ ID NO: 1351-2700. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, \* corresponds to the stop codon.

10

15

20

25

30

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO:1-1350 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO:1-1350. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO:1-1350 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of SEQ ID NO:1-1350.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing

full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

10

15

20

25

30

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO:1 - 1350; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-1350. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing (e.g., SEQ ID NO: 1351-2700); (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO:1-1350; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

10

15

20

25

30

35

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, e.g., in situ hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, butilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and form a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

10

15

20

25

30

35

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases o disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

#### 4. DETAILED DESCRIPTION OF THE INVENTION

#### 4.1 DEFINITIONS

10

15

20

25

30

35

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ

cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived. The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

10

15

20

25

30

35

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can

be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-1350.

5

10

15

20

25

30

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO:1-1350. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4<sup>20</sup> possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match  $(1 \div 4^{25})$  times the increased probability for mismatch at each nucleotide position  $(3 \times 25)$ . The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

5

10

15

20

25

30

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

5

10

15

20

25

30

35

The term "variant" (or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, e.g., recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

5

10

15

20

25

30

35

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, e.g., polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (e.g., microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., E. coli, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

5

10

15

20

25

30

35

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134-143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M NaHPO<sub>4</sub>, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

5

10

15

20

25

30

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (i.e., the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, e.g., mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, e.g., mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (e.g., via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J.

(1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

20

25

30

35

5

10

15

# 4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO:1-1350; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO:1351-2700; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO:1351-2700. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO:1-1350; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 1351-2700. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic

domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

5

10

15

20

25

30

35

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO:1-1350 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO:1-1350 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO:1-1350 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, e.g., at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO:1-1350, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that

are selective for (i.e. specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

5

10

15

20

25

30

35

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO:1-1350, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO:1-1350 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO:1-1350, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic

acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, e.g., by substituting first with conservative choices (e.g., hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (e.g., hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

5

10

15

20

25

30

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., DNA 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, Nucleic Acids Res. 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., Gene 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and Current Protocols in Molecular Biology, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression

Ì

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

5

10

15

20

25

30

35

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO:1-1350, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are

known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

10

15

20

25

30

35

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

# 4.3 ANTISENSE

5

10

15

20

25

30

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1-1350, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID

NO:1351-2700 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO:1-1350 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

10

15

20

Given the coding strand sequences encoding a nucleic acid disclosed herein (e.g., SEQ ID NO:1-1350), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

5

10

15

20

25

30

35

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an -a nomeric nucleic acid molecule. An -a nomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual -units, the strands run parallel to each other (Gaultier et al. (1987) Nucleic Acids Res 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) Nucleic Acids Res 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue et al. (1987) FEBS Lett 215: 327-330).

#### 4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) Nature 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be

designed based upon the nucleotide sequence of a DNA disclosed herein (i.e., SEQ ID NO:1-1350). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, e.g., Cech et al. U.S. Pat. No. 4,987,071; and Cech et al. U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel et al., (1993) Science 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991)

Anticancer Drug Des. 6: 569-84; Helene. et al. (1992) Ann. N.Y. Acad. Sci. 660:27-36; and Maher (1992) Bioassays 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) Bioorg Med Chem 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al. (1996) above; Perry-O'Keefe et al. (1996) PNAS 93: 14670-675.

. 15

20

25

30

35

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup et al. (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn et al. (1996) Nucl Acids Res 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag et al. (1989) Nucl Acid Res 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen et al. (1975) Bioorg Med Chem Lett 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

25

30

35

10

15

20

#### 4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in coamplification of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., Basic Methods in Molecular Biology (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces strains, Candida, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include Escherichia coli, Bacillus subtilis, Salmonella typhimurium, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

5

10

15

20

25

30

35

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

# 4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO:1351-2700 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO:1-1350 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO:1-1350 or (b)

polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO:1351-2700 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO:1351-2700 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO:1351-2700.

5

10

15

20

25

30

35

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

10

15

20

25

30

35

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polypucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, Protein Purification: Principles and Practice, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: A Laboratory Manual; Ausubel et al., Current Protocols in Molecular Biology. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for *e.g.*, small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO:1351-2700.

10

15

20

25

30

35

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, Calif., U.S.A. (the MaxBat<sup>TM</sup> kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

5

10

15

20

25

30

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl<sup>TM</sup> or Cibacrom blue 3GA Sepharose<sup>TM</sup>; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, e.g., targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, e.g., antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

15

20

25

30

35

10

# 4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

### 4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

5

10

15

20

25

30

35

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprises one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers.

Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

#### 4.8 GENE THERAPY

5

10

15

20

25

30

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected ex vivo, in situ, or in vivo by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered in vivo to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

#### 4.9 TRANSGENIC ANIMALS

10

15

20

25

30

35

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

# 4.10 USES AND BIOLOGICAL ACTIVITY

10

15

20

25

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

#### 4.10.1 RESEARCH USES AND UTILITIES

5

10

15

20

25

30

35

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

#### 4.10.2 NUTRITIONAL USES

5

10

15

20

25

30

35

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

# 4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

5

10

15

20

25

30

35

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6-Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11-Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9-Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober.

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095. 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

#### 4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

5

10

20

25

30

35

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells in vivo or ex vivo is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation. manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce 15 large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases: tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

5

10

15

20

25

30

35

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support *e.g.* as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

# 4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

10

15

20

25

30

35

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

15

10

#### 4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

.

20

25

30

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

10

15

20

25

30

35

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine,

kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

20

25

30

5

10

15

# 4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome. autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (e.g., anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastborn et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

10

15

20

25

30

35

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue

transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

10

15

20

25

30

35

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial

immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

10

15

20

25

30

35

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β<sub>2</sub> microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J.

Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

5

- 10

15

20

25

30

35

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

# 4.10.8 ACTIVIN/INHIBIN ACTIVITY

10

15

25

30

35

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

# 4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

5

10

15

20

25

30

35

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

# 4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

### 4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the

invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

5

10

15

20

25

30

35

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine.

Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daumorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen'mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These in vitro models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

#### 4.10.12 RECEPTOR/LIGAND ACTIVITY

15

20

25

30

35

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions

and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14. Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

#### 4.10.13 DRUG SCREENING

10

15

20

25

30

35

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening

utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

5

10

15

20

25

30

35

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science* 282:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., Mol. Biotechnol, 9(3):205-23 (1998); Hruby et al., Curr Opin Chem Biol, 1(1):114-19 (1997); Dorner et al., Bioorg Med Chem, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

# 4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (i.e., increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications *i.e.* phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

5

10

15

20

25

30

### 4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

#### **4.10.16 LEUKEMIAS**

5

10

15

20

35

Leukemias and related disorders may be treated or prevented by administration of a

therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see

Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

### 4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of

therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

- (i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;
- (ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

10

15

20

25

30

- (iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;
- (iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;
- (v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;
- (vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;
- (vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and
  - (viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

5

10

15

20

25

30

35

- (ii) increased sprouting of neurons in culture or in vivo;
- (iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or
  - (iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

#### 4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape);

effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

# 4.10.19 IDENTIFICATION OF POLYMORPHISMS

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or

absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

# 4.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

# 4.11 THERAPEUTIC METHODS

10

15

20

25

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

#### 4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01 µg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1µg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

# 4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

10

15

20

25

30

35

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth

factor (PDGF), transforming growth factors (TGF- $\alpha$  and TGF- $\beta$ ), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

10

15

20

25

30

35

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other

hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

# 4.12.1 ROUTES OF ADMINISTRATION

5

10

15

20

25

30

35

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

#### 4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers

comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

10

15

20

25

30

35

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, tale, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

10

15

20

25

30

35

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral

administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

5

10

20

25

30

35

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other

sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically

acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

10

15

20

25

30

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1 µg to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

10

15

20

25

30

35

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired

patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

20

25

30

10

15

# 4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC<sub>50</sub> as determined in cell culture (i.e., the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD<sub>50</sub> (the dose lethal to 50% of the population) and the ED<sub>50</sub> (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD<sub>50</sub> and ED<sub>50</sub>. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from in vitro data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01  $\mu$ g/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1  $\mu$ g/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

#### 4.12.4 PACKAGING

. 10

15

20

25

30

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

#### 4.13 ANTIBODIES

10

15

20

25

30

35

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain,  $F_{ab}$ ,  $F_{ab}$  and  $F_{(ab)2}$  fragments, and an  $F_{ab}$  expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG<sub>1</sub>, IgG<sub>2</sub>, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, (for example the amino acid sequence shown in SEQ ID NO: 1351), and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, e.g., a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will

indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, e.g., Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

#### 5.13.1 Polyclonal Antibodies

10

15

20

25

30

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A. synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

#### 5.13.2 Monoclonal Antibodies

5

10

15

20

25

30

35

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, <u>Anal. Biochem.</u>, <u>107</u>:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

10

15

20

25

30

35

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium.

Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for

example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

# 5.13.2 Humanized Antibodies

10

15

20

25

30

35

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')2 or other antigenbinding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fy framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

# 5.13.3 Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein.

Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al. (Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse<sup>TM</sup> as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the

immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

#### 5.13.4 Fab Fragment's and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of  $F_{ab}$  expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal  $F_{ab}$  fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an  $F_{(ab)2}$  fragment produced by pepsin digestion of an antibody molecule; (ii) an  $F_{ab}$  fragment generated by reducing the disulfide bridges of an  $F_{(ab)2}$  fragment; (iii) an  $F_{ab}$  fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv)  $F_{v}$  fragments.

#### 5.13.5 Bispecific Antibodies

5

10

15

20

25

30

35

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

5

10

15

20

25

30

35

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., 1991 EMBO J., 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')<sub>2</sub> bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure

wherein intact antibodies are proteolytically cleaved to generate F(ab')<sub>2</sub> fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., <u>J. Exp. Med.</u> 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')<sub>2</sub> molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

10

15

20

25

30

35

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V<sub>H</sub>) connected to a light-chain variable domain (V<sub>L</sub>) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V<sub>H</sub> and V<sub>L</sub> domains of one fragment are forced to pair with the complementary V<sub>L</sub> and V<sub>H</sub> domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., <u>J. Immunol.</u> 147:60 (1991). Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on

a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc R), such as Fc RI (CD64), Fc RII (CD32) and Fc RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

# 5.13.6 Heteroconjugate Antibodies

10

15

20

25

30

35

Heteroconjugate antibodies are also within the scope of the present invention.

Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

# 5.13.7 Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced antitumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

# 5.13.8 Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of

bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include <sup>212</sup>Bi, <sup>131</sup>I, <sup>131</sup>In, <sup>90</sup>Y, and <sup>186</sup>Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

#### 4.14 COMPUTER READABLE SEQUENCES

5

10

15

20

25

30

35

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled

artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

10

15

20

25

30

35

By providing any of the nucleotide sequences SEQ ID NO:1-1350 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO:1-1350 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored

therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

# 4.15 TRIPLE HELIX FORMATION

5

10

15

20

25

30

35

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA.

Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

# 4.16 DIAGNOSTIC ASSAYS AND KITS

10

15

20

25

30

35

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that annual to a polynucleotide of the invention under such conditions, and amplifying annualed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary.

Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization,

amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

#### 4.17 MEDICAL IMAGING

10

15

20

25

30

35

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide in vivo at the target site.

# 4.18 SCREENING ASSAYS

5

10

15

20

25

30

35

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO:1-1350, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

- (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and
- (b) determining whether the agent binds to said protein or said nucleic acid. In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polypucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to

activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

10

15

20

25

30

35

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription

from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

# 10 4.19 USE OF NUCLEIC ACIDS AS PROBES

5

15

20

25

30

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO:1-1350. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO:1-1350 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

# 4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

10

15

20

25

30

Oligonucleotides, i.e., small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata et al., 1985; Dahlen et al., 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller et al., 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, *e.g.*, Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen et al., (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm<sub>7</sub>), is then added to a final concentration of 10 mM 1-MeIm<sub>7</sub>. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

10

15

20

25

30

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-Melm<sub>7</sub>, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

#### 4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

5

10

15

20

25

30

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook et al. (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer et al. (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, CviJI, described by Fitzgerald et al. (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation

of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease CviJI normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme (CviJI\*\*), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald et al. (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a CviJI\*\* digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that CviJI\*\* restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

# 4.22 PREPARATION OF DNA ARRAYS

10

15

20

25

30

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm², depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane.

Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm<sup>2</sup> and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

#### 5.0 EXAMPLES

10

15

20

25

30

#### 5.1 EXAMPLE 1

# Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems

(ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

# 5.2 EXAMPLE 2

# **Novel Contigs**

5

10

20

25

30

The novel contigs of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The sequences for the resulting nucleic acid contigs are designated as SEQ ID NO: 1-1350 and are provided in the attached Sequence Listing. The contigs were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Table 3 sets forth the novel predicted polypeptides (including proteins) encoded by the novel polynucleotides (SEQ ID NO:189-282) of the present invention, and their corresponding nucleotide locations to each of SEQ ID NO: 189-282. Table 3 also indicates the method by which the polypeptide was predicted. Method A refers to a polypeptide obtained by using a software program called FASTY (available from <a href="http://fasta.bioch.virgimia.edu">http://fasta.bioch.virgimia.edu</a>) which selects a polypeptide based on a comparison of the translated novel polynucleotide to known polynucleotides (W.R. Pearson, Methods in Enzymology, 183:63-98 (1990), herein incorporated by reference). Method B refers to a polypeptide obtained by using a software program called GenScan for human/vertebrate sequences (available from Stanford University, Office of Technology Licensing) that predicts the polypeptide based on a probabilistic model of gene structure/compositional properties (C. Burge and S. Karlin, J. Mol. Biol., 268:78-94 (1997), incorporated herein by reference). Method C refers to a polypeptide obtained by using a Hyseq proprietary software program that translates the novel polynucleotide and its complementary strand into six possible amino acid sequences (forward and reverse frames) and chooses the polypeptide with the longest open reading frame.

The nearest neighbor results for SEQ ID NO: 1-1350 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq database October 12, 2000, update 21 (Derwent), using BLAST algorithm. The nearest neighbor result showed the

closest homologue for SEQ ID NO:1-1350. The nearest neighbor results for SEQ ID NO: 1-1350 are shown in Table 2 below.

Tables 1, 2 and 3 follow. Table 1 shows the various tissue sources of SEQ ID NO: 1-1350. Table 2 shows the nearest neighbor result for the assembled contig. The nearest neighbor result shows the closest homolog with an identifiable function for each assemblage. Table 3 contains the start and stop nucleotides for the translated amino acid sequence for which each assemblage encodes. Table 3 also provides a correlation between the amino acid sequences set forth in the Sequence Listing, the nucleotide sequences set forth in the Sequence Listing and the SEQ ID NO. in USSN 09/496,914.

TABLE 1

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
adult brain	GIBCO	AB3001	111 151 188 215 662-665 877 910 927
	1		976 1233 1319
adult brain	GIBCO	ABD003	41 49 74 101 111 120 132 141-142 151
	1		217 225 238 271 317 404 446 469 503
	1		513-514 535 550 564 573 666-669 798
			898 910 927 976 1067 1083 1085 1178
			1254
adult brain	Clontech	ABR001	39 216 238 327 356 535 927 1056 1121
			1178-1180 1199 1251
adult brain	Clontech	ABR006	74 611 949 1034 1136
adult brain	Clontech	ABR008	14 32 41 61 81 86 89 120 132 138 145
			147 188 197 208 225 227-239 250 300-
			303 312 316 328-331 340 357-362 374
	[	•	380 384-391 408 414 446 448 464-467
	,		483 488 495-496 505 512 521 535 550
	}		566 571 577 585 590 594 598 634 641
	Į.		658 666 683 725 742 764 767 786 801
	}	1	805 810 823 826 829 831 836 841 887-
			923 927 934 943 950-951 963 976 995
			1000-1001 1006 1026 1034 1048 1057-
	1		1067 1086 1088 1090 1118 1120 1122-
	. [		1128 1142 1162 1181-1192 1199 1204 1218-1219 1225 1232 1253 1267 1271-
	ĺ		1306 1342 1347 1349-1350
adult brain	Clontech	ABR011	49 238 1219
adult brain	BioChain	ABR012	74 238
adult brain	Invitrogen	ABR013	868 1268
adult brain	Invitrogen	ABT004	49 117 138 191 217 252 291 305 535
want or ann	III VIII OGCII	ADIOUT	566 596 663 670 746 798 816-819 876
			892 898 922 943 963 1034-1036 1121
cultured	Strategene	ADP001	41 74 101 138 211 238 304 537 582
preadipocytes			740 798 883 943 976 1067
adrenal gland	Clontech	ADR002	49 74 101 111 120 127 151 215 238
			240-247 316 330 363-364 404 414 534-
	1		535 833 924-940 950 963 976 1001
			1003 1067-1070 1118 1156 1193-1200
			1325
adult heart	GIBCO	AHR001	38 49 71-72 74-77 79 92 99 101 111
		j	118 129 132 138 151 158-163 182 195-
			203 215 217 238 264 269 353 384 398
			408 434-439 446 504 512-513 519 537
			562-573 577 611-614 616-619 658 661
			671-672 722 734 757-773 815 828-835
	1		874 891 898 919 926-927 976 988
•		· ·	1021 1037 1041 1062 1067 1071 1080
			1083 1093 1122 1131 1185 1201 1254
adult kidney	GMCC	AVDOOL	1308 1331 1335
audit kimity	GIBCO	AKD001	41 49 51 71-74 78-85 94 100-101 103- 107 111 119-120 138 151 157 215 217-
			218 238 250 264 294 304 384 404 440
•	ľ		446 454 477 504-505 509 514 518-519
	1		535 537 564 574-583 620-627 639 653
	ļ		673-675 705 753 789 831 844 851 859
			877 909 918 927 956 963 976 1067
•			1074 1083 1095 1178 1302 1331 1335
adult kidney	Invitrogen	AKT002	11-12 41 49 111-112 215-217 294 316
		1111111	446 487 564 575 844 868 910 927 976
	1	1	1116
adult lung	GIBCO	ALG001	8 101 111 151 187 402 446 490 514
	1	1.2001	

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
		T	518 537 545 549 580 582 592 594 634
			640 651-652 676-678 725 851 873 918
			952 976 1042 1067 1076 1083 1152
lymph node	Clontech	ALN001	8 111 121 151 180-182 188 215 537
•			545 549 651 679-682 789 804-810 868
		İ	873 927 952 976 1042 1059 1335
young liver	GIBCO	ALV001	8 64 79 111 186 215-216 238 446 514
			519 537 564 653 683-684 698 753 798
		1	813 833 840 858 927 976 1038-1039
	•		1051 1085 1224 1245 1256
adult liver	Invitrogen	ALV002	40 71 292-293 305 384 468-469 496
-		1.22.002	505 657 675 714 753 832 844 941-942
			976 1040 1076 1256 1293
adult liver	Clontech	ALV003	976
adult ovary	Invitrogen	AOV001	8 32 36 38 41 49 51 71 74 79-80 101
,	Invice o Ben	110 1001	104 111 120 122-125 138 140 143-149
	<b>\</b>	ł	151 188-190 207-212 215-217 238 264
		1	316 384 409 440 445-446 496 504 512
	·	i e	514 518-519 535 537 549-550 564 566
		1	571 580 582 600 618 638 657 667 681
•		ł	685-697 699 705 722 735-744 761 771
		1	815 833 842-865 868 875-876 918 926-
		j	927 950 952 963 976 1023 1042 1048
		Į	1051 1059 1072 1076 1083 1117 1120
		j	1124 1131 1144 1174 1224 1268 1331
		ł	1335
adult placenta	Clontech	APL001	102 217 238 537 641 700
placenta	Invitrogen	APL002	663 851 1048
adult spleen	GIBCO	ASP001	8 45 74 111 132 140 151 185 217 238
mount opioon	GILCO	751 001	294 414 446 477 504 514 534 545 549
			592 722 873 883 952 976 1041-1042
			1083 1093-1094 1152 1224
testis	GIBCO	ATS001	72 107 111 113 126 140 151 183 215
			238 446 497 537 642 701-706 811 877
·		[	927 962 976 1083 1117 1131
adult bladder	Invitrogen	BLD001	41 151 191 402-405 409 414 496 545
			592 607 706 873 952 1178 1329-1335
bone marrow	Clontech	BMD001	8 58-62 65-68 74 79 108 111 116 137
			147 151 164-174 213-215 238 305-307
		i	374 404 446 460 466 516 519 534 538-
			541 544-546 549-554 566 584 586 592
	.1		596 607 610 628-629 643-645 652 707-
			708 774-789 844 866-871 873 919 927
		·	952 963 976 998 1034 1042 1064 1083
	1	i	1085 1120 1132 1152 1225 1229 1268
			1307 1310
bone marrow	Clontech	BMD002	6 8 37-38 52 74 77 105 111 129 132
		2	210 317 510-511 545 549 581 598 628
	, <b>,</b>		638 724 766 789 844 860 868 873 919
•	l	•	927 952 963 968 976 1042 1111 1141
			1160-1161 1229 1266 1346
bone marrow	Clontech	BMD004	111 238 282 549 1083
adult colon	Invitrogen	CLN001	52 260 264 299 494 536 545 564 592
	m vid og di	OTTION	844 873 877 952 976 1042 1152 1268
	1		1336-1337
adult cervix	BioChain	CVX001	
COUNT CELAIX	DIOCUMIN	CAVOOT	49 51 129 132 151 205 207 238 332-
	j	. 8	335 365-367 392-401 440 466 470-471
	<u> </u>		518 537 597 629 832 877 927 976 1006
	1	1	1085 1117 1129-1134 1192 1202-1205
	1	J	1010 1000 1000
diaphragm	BioChain	DIA002	1219 1309-1328 74 976 1083

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
endothelial cells	Strategene	EDT001	32 40-41 49 74 79 101 111 120 132
	}	1001	138 151 204-206 215-217 238 269 316
·			414 433 505 510 513 550 555 580 582
_	l .		596 675 722 745 798 814 836-841 851
j	}		918 976 1041 1043 1073 1083 1131
	Į.		1331
Genomic clones	Genomic DNA	EPM001	525-532 927
from the short arm	from Genetic	LI MOOI	323-332 321
of chromosome 8	Research		1
Genomic clones	Genomic DNA	EPM003	47 525
from the short arm	from Genetic	121 111003	47 323
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM004	525 927
from the short arm	from Genetic	1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	320 327
of chromosome 8	Research	<u> </u>	1
Genomic clones	Genomic DNA	EPM005	531
from the short arm	from Genetic	1 22 1/2003	331
of chromosome 8	Research		
esophagus	BioChain	ESO002	74 138 238
fetal brain	Clontech	FBR001	441-442 927
fetal brain	Clontech	FBR004	215 893 927 1001
fetal brain	Clontech	FBR006	48 61 101 120 132 138 140 147 208
· · · · · · · · · · · · · · · · · · ·	Cionical	LDKOOO	225 271 317 319 336 359 368 405-414
			519 550 571 594 686 715 722 764 824
			829 836 859 909 927 943 947 963 1057
	1	1	1067-1068 1104 1135-1140 1162 1206-
	)	1	1207 1235 1268 1288 1307-1308 1319
			1338-1350
fetal brain	Clontech	FBRs03	111 446
fetal brain	Invitrogen	FBT002	41 51 120 151 192-194 264 504 512
	ZII VIII OGOII	1 1 1 1 1 1 1	535 683 761 798 820-827 844 876 909
			963 976 1026 1048 1083 1144 1302
fetal heart	Invitrogen	FHR001	446 566 761
fetal kidney	Clontech	FKD001	51 74 111 127 140 151 184 294 537
			550 630-631 1319
fetal kidney	Clontech	FKD002	111 976 1083
fetal kidney	Invitrogen	FKD007	238 974
fetal hmg	Clontech	FLG001	463 566 976 1074 1083 1093
fetal lung	Invitrogen	FLG003	41 238 330 407 415-416 537 573 844
			859 1048 1083 1116 1192
fetal liver-spleen	Columbia	FLS001	8 14 34-35 37 41 43 49 51 54-56 63-64
1	University		69-71 74 77 79 87-90 101 107 110-111
			114 120 128-131 138 140 147 150-155
		ł	197 210 215 217 225 238 312 367 384
	ļ	]	414 440 446 460 468 483 496 504-507
			511-515 518-519 523 533-535 537 541
)			544-545 547-550 555-560 564 566 571
·			577 582 585-586 598 636 646-647 649
ļ	1	·	652 664 698 709-710 714 722-723 731
			735-736 746-753 761 784 798 823 829
	1		832 844 851 858-859 868 873 876 898
	1	•	927 943 949 952 963 976 984 1002
		ì	1021 1023 1040 1042 1044 1050 1083
	1	·	1093 1116 1120 1129 1131 1144 1174
			1217 1251 1254 1256 1302 1308 1311
	1		1319
fetal liver-spleen	Columbia	FLS002	8 36-37 41-46 49 54 64 71 74 79 101
	University		111 120 129 147 207 210 215-216 238
			250 330 353 359 366 383-384 414 478
			505 508-509 511 515-524 534-535 537
			544-545 564 566 571 577 591 598 638
<del></del>		<u> </u>	

Setal liver-spicear	Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
Sil 859 873 876 909 927 949 992 938 984 8002 1023 1024-104 1085 1095 1131 1144 1178 1199 1233 1240-1270 1331 1340     Sil 131 1340		+		663 671 698 714 722 725 727 751 798
Setal liver-spleen		1	[	851 859 873 876 909 927 949 952 983-
131 1144 1178 1199 1233 1240-1276   1331 1340   6etal liver-spleen   Columbia   University   ELSO03   64 535 976 1256   64 535 976 1256   1810 120 138 217 446 468 535 566   1256 1331   6etal liver   Invitrogen   FLV001   S 101 120 138 217 446 468 535 566   1256 1331   6etal liver   Clontech   FLV004   537 926 1256   111 264 312 369-370 404 417-421   425 535 537 577 598 614 836 857 114   1208 1268   1208 1268   1800 1   1208 1268   1800 1   1208 1268   1800 1   1208 1268   1800 1   1208 1268   1800 1   1208 1268   1800 1   1800				
fetal liver         Columbia University         FLS003         64 535 976 1256           fetal liver         Invitrogen         FLV001         8 101 120 138 217 446 468 535 566 580 722 730 749 844 918 943 976 105 1256 1331           fetal liver         Clontech         FLV004         537 926 1256           fetal muscle         Invitrogen         FMS001         51 111 264 312 369-370 404 417-421 425 535 537 577 598 614 836 887 114 1208 1268           fetal muscle         Invitrogen         FMS002         537           fetal skin         Invitrogen         FSK001         13-26 32 41 51 89 107 111 147 151 125 25 64 316 405 422-429 488-494 499 519 534-535 537 506 675 732 859 876 877 889 947 949-950 963 976 1001 1002 1076 1083 1117 1144 1165 1268 1281 1281 12076 1083 1117 1144 1165 1268 1281 1281 1281 1281 1281 128 128 128 1		ļ	Į.	1131 1144 1178 1199 1233 1240-1270
Etal liver			1	1331 1340
Estal liver	fetal liver-spleen	Columbia	FLS003	64 535 976 1256
S80 722 730 749 844 918 943 976 105 1256 1331	•	University		
1256 1331	fetal liver	Invitrogen	FLV001	8 101 120 138 217 446 468 535 566
Fetal liver				580 722 730 749 844 918 943 976 1051
Fetal muscle				1256 1331
fetal muscle Invitrogen FMS002 537 fetal skin Invitrogen FSK001 13-26 32 41 51 89 107 111 147 151 225 264 316 405 422-429 488-494 499 519 534-335 537 566 675 732 859 876 877 898 947 949-950 963 976 1001 1062 1076 1083 1117 1144 1165 1268 1281	fetal liver	Clontech	FLV004	537 926 1256
1208 1268	fetal muscle	Invitrogen	FMS001	
Setal muscle		ì	İ	
FSK001		<u>, , , , , , , , , , , , , , , , , , , </u>		
225 264 316 405 422-429 488-494 496 519 534-535 537 566 675 732 859 876 877 889 897 394-950 963 976 1001 1062 1076 1083 1117 1144 1165 1268 1281				<u> </u>
Signature   Sign	fetal skin	Invitrogen	FSK001	
Setal skin		į.	]	
1062 1076 1083 1117 1144 1165 1268   1281   1282   1282 248-249 301   1316 446 495-503 519 521 534-535 537 522 634 691 877 883 927 944-950 963 976 1001 1075 1142-1143 1171 1218 1243 1308   1243 1308   1449 57 79 87 103 111 120 132-135 138 145 151 188 197 207 215 238 264 271 294 316 367 414 440 446 466 504 271 294 316 367 414 440 446 466 504 271 294 316 367 414 440 446 466 504 271 294 316 367 414 440 446 466 504 271 294 316 367 411 440 446 466 504 271 294 316 367 411 440 446 465 244 271 294 316 367 411 440 446 465 244 271 294 316 367 411 440 446 465 244 271 294 316 367 411 440 446 465 244 271 294 316 367 411 440 446 456 465 204 271 294 316 367 411 440 446 465 244 271 271 272 272 379 832 872 876 883 927 976 1095 1144 1168 1171 1178 121 1335				
Tetal skin				
Fetal skin		i ·		
February   February	F: 1-1-			
BioChain   FUC001   27-33 41 49 151 215 238 248-249 301 316 446 495-503 519 521 534-535 537 582 634 691 877 883 927 944-950 963 976 1001 1075 1142-1143 1171 1218 1243 1308   41 49 57 79 87 103 111 120 132-135 138 145 151 188 197 207 215 238 264 271 294 316 367 414 440 446 466 504 513-514 535 542-543 550 564 571 594 635 648-654 675 711-715 722-723 797 832 872 876 883 927 976 1095 1144 1168 1171 1178 1211 1335				
Signature   Sign	retal spicen			
S82 634 691 877 883 927 944-950 963 976 1001 1075 1142-1143 1171 1218 1243 1308	umbilical cord	BioChain	FUC001	
Proceedings		i		1
fetal brain  GIBCO  HFB001  41 49 57 79 87 103 111 120 132-135 138 145 151 188 197 207 215 238 264 271 294 316 367 414 440 446 466 504 513-514 535 542-543 550 564 571 596 635 648-654 675 711-715 722-723 796 832 872 876 883 927 976 1095 1144 1168 1171 1178 1211 1335  macrophage  Invitrogen  HMP001  238  Infant brain  Columbia  University  B2002  49-50 77 81 89 105 111 136-138 140 151 161 175-179 185 216-217 264 295 299 308-310 371-373 462 476 504 511 513 533 537 564 566 571 655-657 662 683 716-720 723 752 790-803 829 832 838-859 876 898 909 949 976 1045- 1047 1076-1087 1090 1093 1116 1122 1144 1209-1213 1225 1233 1256 1319 1341  infant brain  Columbia  University  Infant brain  Columbia  University  IBS001  S1 111 376 474 790 876 949 1144 120 1221  lung, fibroblast  Strategene  LFB001  Invitrogen  LGT002  1-7 41 74 79 94 115 120 138-139 156 215 217 269 280 296 337 374-375 384 404 446 454 475-480 498 514 518-519 522 537 544 566 4577 9859 868 872-874				
Fetal brain   GIBCO				1
138 145 151 188 197 207 215 238 264 271 294 316 367 414 440 446 466 504 513-514 535 542-543 550 564 571 596 635 648-654 675 711-715 722-723 791 832 872 876 883 927 976 1095 1144 1168 1171 1178 1211 1335	fatal heain	GIRCO	UEDAA1	
271 294 316 367 414 440 446 466 504 513-514 535 542-543 550 564 571 576 635 648-654 675 711-715 722-723 791 832 872 876 883 927 976 1095 1144 1168 1171 1178 1211 1335     macrophage	icini Orani	dibco.	IIIDOOI	
S13-514 535 542-543 550 564 571 596 635 648-654 675 711-715 722-723 794 832 872 876 883 927 976 1095 1144 1168 1171 1178 1211 1335     macrophage		j		
Macrophage				
Macrophage		ļ		635 648-654 675 711-715 722-723 798
Invitrogen		Į		832 872 876 883 927 976 1095 1144
Infant brain				1168 1171 1178 1211 1335
University  151 161 175-179 185 216-217 264 295 299 308-310 371-373 462 476 504 511 513 533 537 564 566 571 655-657 662 683 716-720 723 752 790-803 829 832 858-859 876 898 909 949 976 1045-1047 1076-1087 1090 1093 1116 1122 1144 1209-1213 1225 1233 1256 1319 1341  infant brain  Columbia University  IB2003  IB2003  IB2003  University  IB2003  IB2003  IB2003  University  IB2003  IB2003  University  IB2003  IB2004  IB2004  IB2005  IB2006  IB2006  IB2006  IB2006  IB2007	Invitrogen	HMP001	238	
299 308-310 371-373 462 476 504 511 513 533 537 564 566 571 655-657 662 683 716-720 723 752 790-803 829 832 858-859 876 898 909 949 976 1045- 1047 1076-1087 1090 1093 1116 1122 1144 1209-1213 1225 1233 1256 1319 1341 infant brain Columbia University IB2003 IBM002 IBM002 IBM002 IIBM002 IIBM002 IIBM002 IIBM002 IIBM002 IIBM003 IIBM003 IIBM003 IIBM003 IIBM003 IIBM004 IIBM005 IIBM005 IIBM065 IIBM066 IIII 1376 474 790 876 949 1144 120 IIBM067 IIBM068 IIBM068 IIBM069 IIII 1376 474 790 876 949 1144 120 IIBM069 IIBM069 IIBM069 IIBM069 IIBM069 IIII 1376 474 790 876 949 1144 120 IIBM069 IIBM069 IIBM069 IIBM069 IIII 1376 474 790 876 949 1144 120 IIBM069 IIBM069 IIBM069 IIBM069 IIII 1376 474 790 876 949 1144 120 IIBM069 IIBM069 IIBM069 IIBM069 IIII 1376 474 790 876 949 1144 120 IIBM069 IIBM069 IIBM069 IIII 1376 474 790 876 949 1144 120 IIBM069 IIBM069 IIBM069 IIIII 1376 474 790 876 949 1144 120 IIBM069	infant brain	Columbia	IB2002	49-50 77 81 89 105 111 136-138 140
513 533 537 564 566 571 655-657 662 683 716-720 723 752 790-803 829 832 858-859 876 898 909 949 976 1045- 1047 1076-1087 1090 1093 1116 1122 1144 1209-1213 1225 1233 1256 1319 1341 infant brain Columbia University IB2003 IBM002 IBM002 IBM002 IBM002 IIBM002 IIBM002 IIBM002 IIBM002 IIBM002 IIBM002 IIBM003 IIBM003 IIBM003 IIBM003 IIBM003 IIBM003 IIBM003 IIBM003 IIBM003 IIBM003 III 472-473 753 1214 IUNIVERSITY IUNI		University		151 161 175-179 185 216-217 264 295
683 716-720 723 752 790-803 829 832 858-859 876 898 909 949 976 1045- 1047 1076-1087 1090 1093 1116 1122 1144 1209-1213 1225 1233 1256 1319 1341  infant brain  Columbia University  IB2003  infant brain  Columbia University  IBM002  IBM002  IBM002  IBS001  IBS001  IBS001  IBS001  IBS001  IBS001  ISS001   •			299 308-310 371-373 462 476 504 511-	
858-859 876 898 909 949 976 1045-   1047 1076-1087 1090 1093 1116 1122     1144 1209-1213 1225 1233 1256 1319     1341		`		
1047 1076-1087 1090 1093 1116 1122		,		1
1144 1209-1213 1225 1233 1256 1319   1341     1341			,	5
1341     infant brain		J		I I
infant brain         Columbia University         IB2003         41 50 77 104 132 215 238 508 512-513 519 566 655 714 794 918 943 976 106 1092-1093 1233           infant brain         Columbia University         IBM002         311 472-473 753 1214           infant brain         Columbia University         IBS001         51 111 376 474 790 876 949 1144 120 1221           lung , fibroblast         Strategene         LFB001         151 316 462 514 534 582 675 939 113 120 138-139 156 215 217 269 280 296 337 374-375 384 404 446 454 475-480 498 514 518-519 522 537 545 564 577 597 653 658 705 721-724 754-756 779 859 868 872-874				
University 519 566 655 714 794 918 943 976 106 1092-1093 1233  infant brain Columbia University infant brain Columbia University IBS001 51 111 376 474 790 876 949 1144 120 1221  lung , fibroblast Strategene LFB001 151 316 462 514 534 582 675 939 113 lung tumor Invitrogen LGT002 1-7 41 74 79 94 115 120 138-139 156 215 217 269 280 296 337 374-375 384 404 446 454 475-480 498 514 518-519 522 537 545 564 577 597 653 658 705 721-724 754-756 779 859 868 872-874	infunt benin	Cohumbia	TD2002	
1092-1093 1233     infant brain	mant ciam		נסאססי	1
infant brain Columbia University  infant brain Columbia University  infant brain Columbia University  lung, fibroblast Strategene LFB001 151 316 462 514 534 582 675 939 113  lung tumor Invitrogen LGT002 1-7 41 74 79 94 115 120 138-139 156 215 217 269 280 296 337 374-375 384 404 446 454 475-480 498 514 518-519 522 537 545 564 577 597 653 658 705 721-724 754-756 779 859 868 872-874		Omveisity	•	
University  infant brain  Columbia University  IBS001  Strategene  LFB001  LFB001  LGT002  1-7 41 74 79 94 115 120 138-139 156 215 217 269 280 296 337 374-375 384 404 446 454 475-480 498 514 518-519 522 537 545 564 577 597 653 658 705 721-724 754-756 779 859 868 872-874	infant brain	Columbia	TRM002	
Infant brain         Columbia University         IBS001         51 111 376 474 790 876 949 1144 120 1221           lung, fibroblast         Strategene         LFB001         151 316 462 514 534 582 675 939 113 151 316 462 514 534 582 675 939 113 120 138-139 156 125 217 269 280 296 337 374-375 384 404 446 454 475-480 498 514 518-519 522 537 545 564 577 597 653 658 705 721-724 754-756 779 859 868 872-874	mout viaili		DIVIOUZ	311 412-413 133 1214
University 1221  lung, fibroblast Strategene LFB001 151 316 462 514 534 582 675 939 113  lung tumor Invitrogen LGT002 1-7 41 74 79 94 115 120 138-139 156 215 217 269 280 296 337 374-375 384 404 446 454 475-480 498 514 518-519 522 537 545 564 577 597 653 658 705 721-724 754-756 779 859 868 872-874	infant brain		TRS001	51 111 376 474 790 876 949 1144 1204
lung , fibroblast         Strategene         LFB001         151 316 462 514 534 582 675 939 113           lung tumor         Invitrogen         LGT002         1-7 41 74 79 94 115 120 138-139 156           215 217 269 280 296 337 374-375 384         404 446 454 475-480 498 514 518-519           522 537 545 564 577 597 653 658 705         721-724 754-756 779 859 868 872-874	manus viuill		100001	
lung tumor	hmg fibroblast		LFB001	1
215 217 269 280 296 337 374-375 384 404 446 454 475-480 498 514 518-519 522 537 545 564 577 597 653 658 705 721-724 754-756 779 859 868 872-874				
404 446 454 475-480 498 514 518-519 522 537 545 564 577 597 653 658 705 721-724 754-756 779 859 868 872-874		M. M. M. OSCII	20102	1
522 537 545 564 577 597 653 658 705 721-724 754-756 779 859 868 872-874		1		
721-724 754-756 779 859 868 872-874				l '
1 1 - 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1				
				876-877 919 927 949 951-952 959 976
				1002 1042 1048-1053 1076 1083 1088-
1 1		1		1089 1131 1144-1147 1216-1218 1229

lymphocytes			1000 1011
lymphocytes		.]	1293 1311
1	ATCC	LPC001	41 74 111 132 151 253 316 446 550
		ł	634 844 927 976 1085 1268
leukocyte	GIBCO	LUC001	8 11 41 74 86 91-98 101 109 111 120
			147 151 212 215 218 238 252 288 312-
			314.316 338 359 408 427 443-447 505
			510 512 514 518 534 545 549-550 561
			564 566 571 577 580 582 587-609 615
			632-638 658-659 698 714 725-728 832
			836 841 859 866 873-874 882-883 918-
			919 927 943 952 963 976 1042 1076
			1083 1090 1148 1152 1168 1195 1219-
			1220 1224
leukocyte (	Clontech	LUC003	74 100 215 232 238 339-341 446 545
			657 660 729 873 883 927 952 963 1008
			1042 1116 1120 1149-1150 1215 1222
Melanoma from cell	Clontech	MEL004	210 215 238 342 534 545 592 722 873
line ATCC #CRL			919 929 939 952 976 1071 1118 1218
1424			1235 1245
	Invitrogen	MMG001	8-10 40-41 49 73 80 114 138-140 147
7 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6		1.2.10001	217 250-256 264 297-299 305 377-378
,			398 446 481-486 505 512 537 545 549
			571 592 725 730-733 816 829 836 844
1			868 873 876-877 898 926 943 951-960
			963 976 995 1034 1042 1048 1054-
j		)	1055 1076 1083 1091 1093 1116-1117
			1124 1152 1302
induced neuron cells S	Strategene	NTD001	39 101 111 138 238 361 1225 1251
	Duangono	1112001	1319
retinoid acid induced S	Strategene	NTR001	74 225 976
neuronal cells	Sumogono	1111001	
	Strategene	NTU001	129 225 238 304 313 361 657 976
	Clontech	PIT004	976
	Clontech	PLA003	38 976
	Clontech	PRT001	111 188 238 257-258 564 724 961-966
, , , , , , , , , , , , , , , , , , ,	Olombor.	1111001	1067 1095
rectum	Invitrogen	REC001	238 430-431 841 859 868 963 1001
	шилиодоп	123301	1116
salivary gland (	Clontech	SAL001	8 151 402 432-433 446 496 868 952
Sunvary giants	Cioniccai	SALOUI	976 1083 1120 1151 1184
small intestine (	Clontech	SIN001	8 101 147 215 259-266 446 462 505
Small mostile	Cioniccii	211/001	545 592 660 789 836 866 873 927 952
			963 967-978 1042 1120 1152 1223-
•			1224
skeletal muscle (	Clontech	SKM001	238 302 927 943 992 1031
	Clontech	SPC001	74 111 132 151 215-216 238 264 267-
Spiniar costs	CIOMICCII	SPCOVI	270 343-344 353 379 516 537 566 740
		ļ	828 927 976 979-994 1092 1153-1159
adult spleen (	Clontech	SPLc01	1225 1250
	Clontech	STO001	698 859 1042
2.VIII avi	JOHIEUH	210001	210 238 271-272 537 580 705 918 952
thalamus (	Montrok	TUADO	995 1171
rii anginus (	Clontech	THA002	61 219-220 273-276 312 315 330 596
<del>th</del>	21	7777 4001	963 996-1007 1059 1093 1160-1162
thymus	Clonetech	THM001	8 120 151 208 221 316-317 353 639
			750 867 874 878-881 927 963 1023
			1083 1094-1096 1124
thymus	Clontech	THMc02	8 61 114 129 132 210 225 231 306
ì			317-319 336 340 359 380 398 446 448-
İ			463 512 519 545 554 587 598 698 724-
		i	<b>7</b> 25 789 812 836 868 873 927 947 952

Tissue Origin	RNA Source	Hyseq Library Name	SBQ ID NOS:
			976 1007 1042 1083 1085 1097-1116
			1122 1147 1177 1226-1229 1234 1311
			1313
thyroid gland	Clontech	THR001	14 41 49 76 94 111 144 151 183 188
			210 217 222 253 264 271 277-286 294
			320-326 345-352 361 381-382 446 467
			483 514 534 549-550 564 578 602 649
	1		844 882-883 927 950 956 976 1008-
	1		1028 1076 1083 1117-1120 1142 1163-
			1175 1230-1238 1308
trachea	Clontech	TRC001	223-225 238 287 353-354 514
			545 592 611 873 883-884 927
			952 1029-1031 1042 1151-1152
		1	1170 1176-1177 1239
uterus	Clontech	UTR001	151 226 288-290 355 537 877
		•	885-886 976 1001 1032-1033
	_1		1232

# TABLE 2

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1	B02829	Homo sapiens	Human G protein coupled receptor hRUP5 protein SEQ ID NO:10.	460	100
2	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	111	51
3	R26173	Homo sapiens	Part of Major Yo paraneoplastic antigen (CDR62) encoded by clone pY2.	293	76
4	L29536	Homo sapiens	calcium channel L-type alpha 1 subunit	191	65
5	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	251	50
6	M11507	Homo sapiens	transferrin receptor	120	95
7	AF099100	Homo sapiens	WD-repeat protein 6	1941	93
8	Y92338	Homo sapiens	Human cancer associated antigen precursor from clone NY-RBN-45.	245	82
9	G01343	Homo sapiens	Human secreted protein, SEQ ID NO: 5424.	226	91
10	AJ133798	Homo sapiens	copine VII protein	1127	68
11	G02449	Homo sapiens	Human secreted protein, SEQ ID NO: 6530.	584	99
12	X98330	Homo sapiens	ryanodine receptor 2	282	78
13	AL024498	Homo sapiens	dJ417M14.2 (novel serine/threonine-protein kinase (ortholog of mouse and rat MAK (male germ cell-associated kinase))	293	100
14	AF045577	Pan troglodytes	olfactory receptor OR93Ch	191	36
15	G03131	Homo sapiens	Human secreted protein, SEQ ID NO: 7212.	93	39
16	U26595	Rattus norvegicus	prostaglandin F2a receptor regulatory protein precursor	569	89
17	B08918	Homo sapiens	Human secreted protein sequence encoded by gene 28 SEQ ID NO:75.	99	44
18	Y36203	Homo sapiens	Human secreted protein #75.	165	75
19	U15647	Mus musculus	reverse transcriptase	106	40
20	G02701	Homo sapiens	Human secreted protein, SEQ ID NO: 6782.	544	100
21	Y35923	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 172.	1691	100
22	G04030	Homo sapiens	Human secreted protein, SEQ ID NO: 8111.	380	96
23	G02455	Homo sapiens	Human secreted protein, SEQ ID NO: 6536.	123	50
24	AF036329	Homo sapiens	gonadotropin-releasing hormone precursor, second form	284	90
25	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	96	32
26	S80119	Rattus sp.	reverse transcriptase homolog	100	34
27	U83303	Homo sapiens	line-1 reverse transcriptase	101	35
28	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	135	45

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:	604067	77	The second of th	Score 83	42
29	G04067 G02872	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 8148.  Human secreted protein, SEQ ID NO: 6953.	116	72
30 31	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	96	67
32	G03224	Homo sapiens	Human secreted protein, SEQ ID NO: 7305.	58	32
33	Y66688	Homo sapiens	Membrane-bound protein PRO1152.	2457	98
34	Y87071	Homo sapiens	Human secreted protein sequence SEQ ID	348	95
<b>.</b>	10,0,1	Tional Suprema	NO:110.		
35	U15131	Homo sapiens	p126	182	48
36	Y73464	Homo sapiens	Human secreted protein clone yl4_1 protein sequence SEQ ID NO:150.	982	90 .
37	AL133215	Homo sapiens	bA108L7.6 (semaphorin 4G (sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain))	687	99
38	AC067969	amino acids 3338-4088	Homo sapiens ryanodine receptor 1 (skeletal)	386	66
39	AL031588	Homo sapiens	dJ1163J1.1 (mostly supported by GENSCAN, FGENES and GENEWISE)	493	76
40	G03628	Homo sapiens	Human secreted protein, SEQ ID NO: 7709.	110	51
41	AF132969	Horno sapiens	CG1-35 protein	228	68
42	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	220	88
43	X61048	Hydra sp.	mini-collagen	105	35
44	M76546	Helianthus annuus	hydroxyproline-rich protein	110	31
45	U82288	Caenorhabditi s elegans	Rac-like GTPase	139	70
46	G03477	Homo sapiens	Human secreted protein, SEQ ID NO: 7558.	118	58
47	AF090942	Homo sapiens	PRO0657	113	63
48	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	90	59
49	AJ005560	Mus musculus	SPR2B protein	72	56
50	G02450	Homo sapiens	Human secreted protein, SEQ ID NO: 6531.	385	98
51	Y91649	Homo sapiens	Human secreted protein sequence encoded by gene 60 SEQ ID NO:322.	973	94
52	U93563	Homo sapiens	putative p150	105	38
53	Y55927	Homo sapiens	Human STLK2 protein.	699	85
54	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	145	56 74
55	AB008175	Mus musculus	hepatic nuclear factor 1-beta short form	356	
56	M68941	Homo sapiens	protein-tyrosine phophatase	165	41
57	AL031600	Homo sapiens	c390E6.1 (chloride channel 7)	338	76
58	AF011417	Mus musculus	putative pheromone receptor	143	55
59	AF167320	Mus musculus	zinc finger protein ZFP113	558	68
60	U73036	Homo sapicns	interferon regultory factor 7	263	96
61	X07984	Mus musculus	protein-tyrosine kinase	297	69
62	Y29861	Homo sapiens	Human secreted protein clone cb98_4.	791	98
63 64	U35376 AF265555	Homo sapiens Homo sapiens	repressor transcriptional factor ubiquitin-conjugating BIR-domain enzyme	485 785	65 74
			APOLLON	1	10-
65	G03883	Homo sapiens	Human secreted protein, SEQ ID NO: 7964.	88	95
66	AF177390	Manduca sexta	antennal specific membrane protein AMP	274	54
67	AB040800	Homo sapiens	SREB2	614	100
68	AF030027	Equine herpesvirus 4	24	213	26
69	G02965	Homo sapiens	Human secreted protein, SEQ ID NO: 7046.	261	95
70	W75770	Homo sapiens	Human oxidoreductase YTFO3.	1144	98
71	AB011135	Homo sapiens	KIAA0563 protein	239	76
72	AB014885	Halocynthia roretzi	HrPOPK-1	813	78
73	AF045454	Cavia porcellus	phospholipase B	955	73
74	J02870	Mus	laminin receptor ·	308	61

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:	<del> </del>	-		Score	<u> </u>
75	Y00826	Rattus			
15	100826	norvegicus	gp210 (AA 1-1886)	413	84
76	AF117754	Homo sapiens	thoroid beautiful and a second	251	-
70	AT11//34	nomo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	351	54
77	Y38422	Homo sapiens	Human secreted protein.	469	75
78	Y14596	Homo sapiens	Human T-type voltage-gated Ca channel alpha-	468	76 99
70	114390	rionio sapiens	1-I (hCavT3).	1357	99
79	Y14591	Human	APM-1 protein	767	100
• •	1	papillomaviru	1 mm i process	1 "0"	100
		s type 68		1	
80	AL137802	Homo sapiens	dJ798A10.2 (KIAA0445 protein)	71	34
81	AP000383	Arabidopsis	protein arginine N-methyltransferase-like protein	359	65
		thaliana	proved and it moust it amount into protein	1 337	"
82	L46815	Mus	DNA binding protein Rc	895	75
		musculus			1 "
83	G01600	Homo sapiens	Human secreted protein, SEQ ID NO: 5681.	315	96
84	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	538	71
			designated HSCOP-6.	}	1
85	AB029002	Homo sapiens	KIAA1079 protein	134	42
86	Y28678	Homo sapiens	Human cw272_7 secreted protein.	325	62
87	Y99368	Homo sapiens	Human PRO1326 (UNQ686) amino acid	156	48
		<u> </u>	sequence SEQ ID NO:100.		J
88	AJ225124	Mus	hyperpolarization-activated cation channel,	487	95
		musculus	HAC3		
89	AF177203	Homo sapiens	cerebral cell adhesion molecule	290	56
90	Y28280	Homo sapiens	Human G-protein coupled receptor GRIR-2.	326	79
91	L39891	Homo sapiens	polycystic kidney disease-associated protein	1751	95
92 93	AF064876 AF170723	Homo sapiens	ion channel BCNG-1	953	99
93	X13292	Homo sapiens	protein kinase STK10	401	53
<del>"</del>	X13292	Trypanosoma brucei	GPI-phospholipase C (AA 1 - 358)	151	37
95	Y34127	Homo sapiens	Human potassium channel K+Hnovl 1.	661	99
96	X03638	Rattus	sodium channel protein I (aa 1-2009)	1775	92
<b>J</b>	705050	norvegicus	Sommit chariter protein 1 (az 1-2003)	1773	32
97	AF134213	Homo sapiens	ubiquitin-specific protease	1995	99
98	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	213	38
99	AF021935	Rattus	mytonic dystrophy kinase-related Cdc42-binding	675	48
		norvegicus	kinase		"
100	AF279265	Homo sapiens	putative anion transporter 1	867	98
101	AC007878	Homo sapiens	match to nuclear protein, NP220; note: sequence	160	60
	<u> </u>	1	difference at residue 58		
102	U22829	Mus	P2Y purinoceptor	264	42
		musculus	×		
103	Y45023	Homo sapiens	Human sensory transduction G-protein coupled	516	99
104	7/04/022	ļ.,	receptor-B3.		
104	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	787	98
105	Y87342	Homo sapiens	Human signal peptide containing protein HSPP-	343	57
102	AE160210	177	119 SEQ ID NO:119.		<u> </u>
106	AF169312	Homo sapiens	hepatic angiopoietin-related protein	212	67
107	AF116657	Homo sapiens	PRO1310	74	52
108	AE000401	Escherichia	sialic acid transporter	587	96
109	Y38395	Homo sapiens	Human consider a state and discovery	702	100
110	Y78801	Homo sapiens	Human secreted protein encoded by gene No. 10.	693	100
110	1 /00/1	Aomo Sapiens	Hydrophobic domain containing protein clone HP00631 amino acid sequence.	182	94
111	Z25535	Homo sapiens	nuclear pore complex protein hnup153	164	05
112	Y94939	Homo sapiens	Human secreted protein clone ye90 1 protein	464	85
112	1 24737	Tronio sabicus	sequence SEQ ID NO:84.	274	51
113	AF016365	Homo sapiens	hexokinase 1 isoform td	301	71
114	AC007956	Homo sapiens	unknown		71
115	M83738	Homo sapiens	protein-tyrosine phosphatase	520 251	75
116	AL157952	Homo sapiens	dJ875K15.1.1 (ets homologous factor (ets-	484	92
-10	AL 131732	Tromo sapiens	domain transcription factor ESE-3A, isoform 1))	70 <del>4</del>	91
	1		Human Aurora-2.		I

SEQ	Accession	Species	Description	Smith-	%
ID	No.	Į -		Waterman	Identity
NO:		<u> </u>	·	Score	1
118	L41816	Homo sapiens	cam kinase I	407	62
119	AJ006710	Rattus	phosphatidylinositol 3-kinase	627	93
		norvegicus	_		
120	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor, PDPr	1646	94
121	S39392	Homo sapiens	protein tyrosine phosphatase, PTPase {EC 3.1.3.48}	373	68
122	U60805	Homo sapiens	oncostatin-M specific receptor beta subunit	262	88
123	Y44403	Homo sapiens	Human truncated tankyrase-1.	111	35
124	U88167	Caenorhabditi s elegans	contains similarity to C2 domains	219	29
125	AF300648	Homo sapiens	guanine nucleotide binding protein beta subunit	693	90
126	AB021861	Mus musculus	apoptosis signal-regulating kinase 2	153	65
127	AF305210	Homo sapiens	concentrative Na+-nucleoside cotransporter hCNT3	807	97
128	M90360	Homo sapiens	protein kinase	220	73
129	D32202	Homo sapiens	alpha 1C adrenergic receptor isoform 2	574	86
130	AF208043	Homo sapiens	IF116b	496	67
131	AF201734	Mus musculus	testis specific serine kinase-3	800	87
132	AF112886	Bos taurus	differentiation enhancing factor 1	159	74
133	AJ278314	Homo sapiens	phospholipase C-beta-1b	554	85
134	W74802	Homo sapiens	Human secreted protein encoded by gene 73	1157	87
			clone HSQEL25.	1137	"
135	AB020335	Homo sapiens	Pancreas-specific gene	668	96
136	W80408	Homo sapiens	A secreted protein encoded by clone dt674 2.	866	98
137	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95% similarity to P49205 (PID:g1345860)	5041	99
138	Y96736	Homo sapiens	PRO3434, a novel secreted protein.	891	100
139	AB024034	Arabidopsis thaliana	DNA-damage inducible protein DDI1-like	147	55
140	W97809	Homo sapiens	Human GTPase regulator GRAF.	248	56
141	Y51557	Homo sapiens	Human PLA2 protein.	125	46
142	AF090113	Rattus norvegicus	AMPA receptor binding protein	623	93
143	W26642	Homo sapiens	Human RECK cancer-inhibiting protein.	641	82
144	U87306	Rattus norvegicus	transmembrane receptor UNC5H2	578	84
145	AF264014	Homo sapiens	scavenger receptor cysteine-rich type I protein M160 precursor	727	92
146	W63683	Homo sapiens	Human secreted protein 3.	140	40
147	M96264	Homo sapiens	galactose-1-phosphate uridyl transferase	513	81
148	D64014	Escherichia coli	HrsA	818	90
149	M83316	Escherichia coli	pppGpp phosphohydrolase	915	95
150	AL163279	Homo sapiens	homolog to cAMP response element binding and beta transducin family proteins	1261	99
151	AF179867	Homo sapiens	STE20-like kinase	940	99
152	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain ligand (clone 3TW).	392	61
153	AF151859	Homo sapiens	CGI-101 protein	370	92
154	X66957	Homo sapiens	hexokinase type 1	489	81
155	Y16355	Homo sapiens	alternatively spliced form	432	92
156	G00857	Homo sapiens	Human secreted protein, SEQ ID NO: 4938.	349	78
157	AF159455	Mus musculus	zinc finger protein	352	74
158	L76191	Homo sapiens	interleukin-1 receptor-associated kinase	537	76
159	AP001743	Homo sapiens	putative gene, ankirin like, possible dual specifity Ser/Thr/Tyr kinase domain	670	98
160	AJ250425	Rattus norvegicus	Collybistin 1	556	74
	G02885	Homo sapiens	Human secreted protein, SEQ ID NO: 6966.	370	100

SEQ	Accession	Species	Description	Smith-	1%
D`	No.	1 2		Waterman	Identity
NO:				Score	
162	Z22968	Homo sapiens	M130 antigen	610	100
163	AF181121	Homo sapiens	ATP-dependent Ca2+ pump PMR1	336	92
164	AF055636	Homo sapiens	leucine-rich glioma-inactivated protein precursor	455	94
165	AF160798	Rattus	calcium transporter CaT1	700	96
		norvegicus	•		
166	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	327	45
167	Y48607	Homo sapiens	Human breast tumour-associated protein 68.	1072	99
168	AB020741	Mus musculus	NIK-related kinase	197	43
169	AF252293	Homo sapiens	PAR3	596	44
170	U59429	Cricetinae	diacylglycerol kinase eta	481	82
		gen. sp.			<u> </u>
171	AF035268	Homo sapiens	phosphatidylserine-specific phospholipase A1	386	42
172	AF127085	Mus musculus	semaphorin cytoplasmic domain-associated protein 3B	507	82
173	Y27918	Homo sapiens	Human secreted protein encoded by gene No.	653	99
174	G02979	Vice	123.	520	100
174 175	U36488	Homo sapiens Mus	Human secreted protein, SEQ ID NO: 7060.	538	97
		musculus	embryonic stem cell phosphatase	168	55 .
176	W95629	Homo sapiens	Homo sapiens secreted protein gene clone gm196 4.	1022	100
177	AF289023	Homo sapiens	formiminotransferase cyclodeaminase form D	255	93
178	X04936	Homo sapiens	T-cell receptor alpha-chain (413 is 2nd base in	710	99
			codon)		
179	AF127481	Homo sapiens	non-ocogenic Rho GTPase-specific GTP exchange factor	175	80
180	G00978	Homo sapiens	Human secreted protein, SEQ ID NO: 5059.	517	94
181	Y66645	Homo sapiens.	Membrane-bound protein PRO1310.	671	96
182	AF110640	Homo sapiens	orphan seven-transmembrane receptor	862	100
183	AB020854	Bos taurus	orphan transporter short splicing variant	766	84
184	AF169691	Homo sapiens	cadherin-like protein VR8	375	38
185	AF126372	Homo sapiens	thyrotropin-releasing hormone degrading ectoenzyme	985	99
186	L20966	Homo sapiens	phosphodiesterase	541	76
187	G02920	Homo sapiens	Human secreted protein, SEQ ID NO: 7001.	254	93
188	Y94918	Homo sapiens	Human secreted protein clone dd504_18 protein	301	98
189	Y66713	Homo sapiens	sequence SEQ ID NO:42.	694	100
			Membrane-bound protein PRO1309.		
190 191	G03244 U36771	Homo sapiens Rattus	Human secreted protein, SEQ ID NO: 7325. sn-glycerol 3-phosphate acyltransferase	331 707	73 92
100		norvegicus	S	1.50	-
192	R05935	Homo sapiens	Secreted GPIIb subunit of multiple subunit polypeptide (MSP)GPIIb-IIIa.	157	72
193	M92084	Theileria parva	casein kinase II alpha subunit	364	50
194	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	448	90
195	W95631	Homo sapiens	Homo sapiens secreted protein gene clone	382	49
196	AF255614	Rattus	hj968_2. scaffolding protein SLIPR	680	99
197	AC021640	norvegicus  Arabidopsis	putative phosphatidate phosphohydrolase	300	41
		thaliana			
198	AF073967	Mus musculus domesticus	olfactory receptor	316	43
199	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	617	98
	AF117948	Homo sapiens	pancreas-cariched phospholipase C	625	89
200		Homo sapiens	CDC42-binding protein kinase beta	636	94
	AF128625	1 TIONIO SONICHA			
201	AF128625 AF117946		Link quanine nucleotide exchange factor II	1303	100
200 201 202 203	AF128625 AF117946 Y53021	Homo sapiens Homo sapiens	Link guanine nucleotide exchange factor II  Human secreted protein clone qc646_1 protein	1303 701	100 99
201 202	AF117946	Homo sapiens			

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:				Score	
-006	7110015	<del> </del>	{ovarian cancer critical region of deletion}		<u> </u>
206	U18315 AF255342	Sus scrofa	parathyroid receptor	122	60
207	S52051	Homo sapiens	putative pheromone receptor V1RL1 long form	170	96
209	W63683	Rattus sp. Homo sapiens	neurotransmitter transporter Human secreted protein 3.	715 840	94
210	D79992	Homo sapiens	similar to Drosophila photoreceptor cell-specific	541	82
			protein, calphotin.		
211	AF117948	Homo sapiens	pancreas-enriched phospholipase C	1348	99
212	U81035	Rattus norvegicus	ankyrin binding cell adhesion molecule neurofascin	471	69
213	AF154846	Homo sapiens	zinc finger protein	798	56
214	AF102777	Mus musculus	FYVE finger-containing phosphoinositide kinase	933	93
215	AL163303	Homo sapiens	putative gene containing transmembrane domain	523	89
216	U26595	Rattus norvegicus	prostaglandin F2a receptor regulatory protein precursor	563	78
217	G04095	Homo sapiens	Human secreted protein, SEQ ID NO: 8176.	644	98
218	X75756	Homo sapiens	protein kinase C mu	314	81
219	Y66723	Homo sapiens	Membrane-bound protein PRO1100.	770	98
220	D88577	Mus musculus	Kupffer cell receptor	567	40
221	AF258465	Homo sapiens	OTRPC4	853	100
222	AF021935	Rattus norvegicus	mytonic dystrophy kinase-related Cdc42-binding kinase	636	96
223	AL136527	Homo sapiens	bA215B13.1 (A kinase (PRKA) anchor protein	693	100
224	AB032417	Homo sapiens	WNT receptor Frizzled-4	690	99
225	AF030430	Mus	semaphorin VIa	703	68
		musculus	,		
226	AE000218	Escherichia coli	putative dihydroxyacetone kinase (EC 2.7.1.2)	297	39
227	AF302150	Homo sapiens	phosphoinositol 3-phosphate-binding protein-2	2080	100
228	AB024573	Mus musculus	GTP-binding like protein 2	265	88
229	AF122924	Xenopus laevis	Wnt inhibitory factor-1	316	40
230	G03205	Homo sapiens	Human secreted protein, SEQ ID NO: 7286.	229	100
231	X98260	Homo sapiens	M-phase phosphoprotein 11	265	92
232	R92754	Homo sapiens	Human growth differentiation factor-12.	682	95
233	R75111	Homo sapiens	Glycosyl-phosphatidylinositol-specific phospholipase-D.	290	100
234	W69431	Homo sapiens	Human secreted protein cw1233 3.	235	97
235	Y08686	Homo sapiens	serine palmitoyltransferase, subunit II	859	81
236	AF118275	Homo sapiens	atrophin-related protein ARP	117	37
237	X81466	Mus musculus	Embryo Brain Kinase	460	62
238	U64857	Caenorhabditi s elegans	similar to the BPTI/Kunitz family of inhibitors; most similar to tissue factor pathway inhibitor precursor (TFPI)	284	33
239	AJ250840	Mus musculus	serine/threonine protein kinase	739	63
240	AJ223472	Mus musculus	transcription elongation factor TFIIS.h	222	38
241	Y94906	Homo sapiens	Human secreted protein clone rb649_3 protein	353	52
242	AF169301	Homo sapiens	sequence SEQ ID NO:18. Na+/sulfate cotransporter SUT-1	591	99
243	L22022	Rattus	orphan transporter v7-3	667	93
244	AF016191	norvegicus Rattus	potassium channel	1043	98
246	A E0003055	norvegicus	4		
245	AF097366	Homo sapiens	cone sodium-calcium potassium exchanger	645	98
246 247	Y29868 AF180475	Homo sapiens	Human secreted protein clone pp325_9.	497	98
248	Y17227	Homo sapiens Homo sapiens	Not4-Np Human secreted protein (clone ya1-1).	188 690	83
		I LIULUU SHURGIIS	Lighton Scatter Divient (Clone val-1).	リプリ	99

SEQ ID	Accession	Species	Description	Smith-	1%
	No.	1 - 7 - 11 - 1		Waterman	Identity
NO:	1	1		Score	rectition
110.	<del> </del>	sexta	protein SCLP	00010	<del> </del>
250	AF192756	Kaposi's	Orf73	134	34
230	AF 192130	sarcoma-	01173	154	34
	1	associated	1	1	ì
				Į	1
***	<del> </del>	herpesvirus	<u> </u>	<u> </u>	<del></del>
251	AB022694	Homo sapiens	MOK protein kinase	209	83
252	W55045	Homo sapiens	Neural adhesion molecule (ethb0018f2 product).	469	100
253	L46815	Mus	DNA binding protein Rc	251	67
	İ	musculus			
254	W68505	Homo sapiens	Human acid sensing ionic channel.	173	82
255	AF070066	Mus	Citron-K kinase	1201	98
	1	musculus		120.	1 20
256	G02491	Homo sapiens	Human secreted protein, SEQ ID NO: 6572.	460	100
257	Z12841			1	
237	Z12841	Oryctolagus	Phospholipase	368	80
	1-22-2-	cuniculus		<del> </del>	
258	Y95436	Homo sapiens	Human calcium channel SOC-3/CRAC-2.	1857	99
259	AJ222968	Mus	L-periaxin	430	72
	<u>]</u>	musculus		1	1.
260	AJ250839	Homo sapiens	serine/threonine protein kinase	861	100
261	AJ249977	Homo sapiens	AMP-activated protein kinase gamma 3 subunit	758	98
262	AF141386	Rattus	SLIT-2	198	40
	1	norvegicus		1	"
263	AF022859	Homo sapiens	neuropilin-2(a0)	335	62
264	AF160477	Homo sapiens	Ig superfamily receptor LNIR precursor	387	91
	Y44662				
265	Y44002	Homo sapiens	Human 14273 G-protein coupled receptor	636	99
	<del> </del>	<del> </del>	(GPCR).		<u> </u>
266	U27269	Mus	sodium glucose cotransporter	204	56
	l	musculus			
267	AF124491	Homo sapiens	ARF GTPase-activating protein GIT2	159	75
268	AF127389	Rattus	putative taste receptor TR1	209	39
	1	norvegicus	•		Ī
269	X98296	Homo sapiens	ubiquitin hydrolase	215	95
270	X78482	Streptococcus	Fc-gamma receptor	129	26
	1	pyogenes			1
				<u> </u>	
271	AB009883	Nicotiana	I KED	109	26
271	AB009883	Nicotiana tahacum	KED	109	26
_,_		tabacum			_
271	AB009883 AF137367	tabacum Mus	VPS10 domain receptor protein SORCS	109 899	26 97
272	AF137367	tabacum Mus musculus	VPS10 domain receptor protein SORCS	899	97
272		tabacum Mus musculus Rattus			_
272 273	AF137367 L34938	tabacum Mus musculus Rattus norvegicus	VPS10 domain receptor protein SORCS	899	97
272 273	AF137367	tabacum Mus musculus Rattus	VPS10 domain receptor protein SORCS Ionotropic glutamate receptor dJ413H6.1.1 (hamster Androgen-dependent	899	97
272 273	AF137367 L34938	tabacum Mus musculus Rattus norvegicus	VPS10 domain receptor protein SORCS lonotropic glutamate receptor dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein)	899	97
272 273 274	AF137367 L34938	tabacum Mus musculus Rattus norvegicus Homo sapiens	VPS10 domain receptor protein SORCS lonotropic glutamate receptor dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1)	899	97
272 273 274	AF137367 L34938	tabacum Mus musculus Rattus norvegicus	VPS10 domain receptor protein SORCS  lonotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1)  ubiquitin-conjugating BIR-domain enzyme	899	97
272 273 274	AF137367 L34938 AL022724	tabacum Mus musculus Rattus norvegicus Homo sapiens	VPS10 domain receptor protein SORCS ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON	899 460 188	97 86 74
272 273 274 275	AF137367 L34938 AL022724	tabacum Mus musculus Rattus norvegicus Homo sapiens	VPS10 domain receptor protein SORCS ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON	899 460 188	97 86 74
272 273 274 275 276	AF137367 L34938 AL022724 AF265555 G02872	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens	VPS10 domain receptor protein SORCS ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953.	899 460 188 173	97 86 74 94
272 273 274 275 276 277	AF137367 L34938 AL022724 AF265555 G02872 L40380	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens	VPS10 domain receptor protein SORCS lonotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor	899 460 188 173 148 430	97 86 74 94 56 61
272 273 274 275 276 277 278	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens	VPS10 domain receptor protein SORCS ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein	899 460 188 173 148 430 283	97 86 74 94 56 61 96
272 273 274 275 276 277 278	AF137367 L34938 AL022724 AF265555 G02872 L40380	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Arabidopsis	VPS10 domain receptor protein SORCS ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953, thyroid receptor interactor KIAA1631 protein Contains PF100069 Eukaryotic protein kinase	899 460 188 173 148 430	97 86 74 94 56 61
272 273 274 275 276 277 278 279	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens thomo sapiens homo sapiens	VPS10 domain receptor protein SORCS ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein Contains PF[00069 Eukaryotic protein kinase domain.	899 460 188 173 148 430 283 157	97 86 74 94 56 61 96 43
272 273 274 275 276 277 278 279	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	VPS10 domain receptor protein SORCS ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein Contains PF[00069 Eukaryotic protein kinase domain. protein-tyrosine phosphatase	899 460 188 173 148 430 283 157	97 86 74 94 56 61 96 43
272 273 274 275 276 277 278 279 280 281	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738  AK024397	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Arabidopsis thaliana Homo sapiens	VPS10 domain receptor protein SORCS ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein Contains PF 00069 Eukaryotic protein kinase domain. protein-tyrosine phosphatase unnamed protein product	899 460 188 173 148 430 283 157 181 439	97 86 74 94 56 61 96 43 73 91
272 273 274 275 276 277 278 279 280 281 282	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738  AK024397  AF141326	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	VPS10 domain receptor protein SORCS  ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1)  ubiquitin-conjugating BIR-domain enzyme APOLLON  Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor  KIAA1631 protein  Contains PF 00069 Eukaryotic protein kinase domain.  protein-tyrosine phosphatase unnamed protein product  RNA helicase HDB/DICE1	899 460 188 173 148 430 283 157	97 86 74 94 56 61 96 43 73 91 84
272 273 274 275 276 277 278 279 280 281 282	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738  AK024397	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Arabidopsis thaliana Homo sapiens	VPS10 domain receptor protein SORCS ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein Contains PF 00069 Eukaryotic protein kinase domain. protein-tyrosine phosphatase unnamed protein product	899 460 188 173 148 430 283 157 181 439	97 86 74 94 56 61 96 43 73 91
272 273 274 275 276 277 278 279 280 281 282	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738  AK024397  AF141326	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Arabidopsis thaliana Homo sapiens Homo sapiens	VPS10 domain receptor protein SORCS  ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1)  ubiquitin-conjugating BIR-domain enzyme APOLLON  Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor  KIAA1631 protein  Contains PF 00069 Eukaryotic protein kinase domain.  protein-tyrosine phosphatase unnamed protein product  RNA helicase HDB/DICE1	899 460 188 173 148 430 283 157 181 439 497	97 86 74 94 56 61 96 43 73 91 84
272 273 274 275 276 277 278 279 280 281 282 283	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738  AK024397  AF141326  AF156530	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Arabidopsis thaliana Homo sapiens Homo sapiens Mus musculus	VPS10 domain receptor protein SORCS  ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1)  ubiquitin-conjugating BIR-domain enzyme APOLLON  Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor  KIAA1631 protein  Contains PF]00069 Eukaryotic protein kinase domain.  protein-tyrosine phosphatase unnamed protein product  RNA helicase HDB/DICE1  ETS-domain transcriptional repressor PE1	899  460  188  173  148  430  283  157  181  439  497  605	97 86 74 94 56 61 96 43 73 91 84 76
272 273 274 275 276 277 278 279 280 281 282 283	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738  AK024397  AF141326	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Arabidopsis thaliana Homo sapiens Homo sapiens Homo sapiens	VPS10 domain receptor protein SORCS ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein Contains PF 00069 Eukaryotic protein kinase domain. protein-tyrosine phosphatase unnamed protein product RNA helicase HDB/DICE1 ETS-domain transcriptional repressor PE1 Human secreted protein clone cs756_2 alternate	899 460 188 173 148 430 283 157 181 439 497	97 86 74 94 56 61 96 43 73 91 84
2772 2773 2774 2775 2776 2777 2778 2779 280 281 282 283 284	AF137367 L34938 AL022724 AF265555 G02872 L40380 AB046851 AC008075 M83738 AK024397 AF141326 AF156530 Y29336	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Arabidopsis thaliana Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus Homo sapiens	VPS10 domain receptor protein SORCS  lonotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATTVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON  Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein Contains PF100069 Eukaryotic protein kinase domain. protein-tyrosine phosphatase unnamed protein product RNA helicase HDB/DICE1 ETS-domain transcriptional repressor PE1  Human secreted protein clone cs756_2 alternate reading frame protein.	899 460 188 173 148 430 283 157 181 439 497 605 647	97 86 74 94 56 61 96 43 73 91 84 76
2772 2773 2774 2775 2776 2777 2778 2779 280 281 282 283 284	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738  AK024397  AF141326  AF156530	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Arabidopsis thaliana Homo sapiens Homo sapiens Mus musculus	VPS10 domain receptor protein SORCS  lonotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATTVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON  Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein Contains PF100069 Eukaryotic protein kinase domain. protein-tyrosine phosphatase unnamed protein product RNA helicase HDB/DICE1 ETS-domain transcriptional repressor PE1  Human secreted protein clone cs756_2 alternate reading frame protein. Human secreted protein clone yc25_1 protein	899  460  188  173  148  430  283  157  181  439  497  605	97 86 74 94 56 61 96 43 73 91 84 76
272 273 274 275 276 277 278 279 280 281 282 283 284 285	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738  AK024397  AF141326  AF156530  Y29336  Y73402	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Arabidopsis thaliana Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus Homo sapiens	VPS10 domain receptor protein SORCS lonotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein Contains PF100069 Eukaryotic protein kinase domain. protein-tyrosine phosphatase unnamed protein product RNA helicase HDB/DICE1 ETS-domain transcriptional repressor PE1 Human secreted protein clone cs756_2 alternate reading frame protein. Human secreted protein clone yc25_1 protein sequence SEQ ID NO:26.	899 460 188 173 148 430 283 157 181 439 497 605 647	97 86 74 94 56 61 96 43 73 91 84 76
272 273 274 275 276 277 278 279 280 281 282 283 284 285	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738  AK024397  AF141326  AF156530  Y29336  Y73402  AF016411	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Arabidopsis thaliana Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus Homo sapiens Homo sapiens	VPS10 domain receptor protein SORCS lonotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein Contains PF100069 Eukaryotic protein kinase domain. protein-tyrosine phosphatase unnamed protein product RNA helicase HDB/DICE1 ETS-domain transcriptional repressor PE1 Human secreted protein clone cs756_2 alternate reading frame protein. Human secreted protein clone yc25_1 protein sequence SEQ ID NO:26. KCNA3.1B	899 460 188 173 148 430 283 157 181 439 497 605 647 300	97 86 74 94 56 61 96 43 73 91 84 76 100
272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738  AK024397  AF141326  AF156530  Y29336  Y73402  AF016411  W89253	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Arabidopsis thaliana Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus Homo sapiens	VPS10 domain receptor protein SORCS lonotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein Contains PF100069 Eukaryotic protein kinase domain. protein-tyrosine phosphatase unnamed protein product RNA helicase HDB/DICE1 ETS-domain transcriptional repressor PE1 Human secreted protein clone cs756_2 alternate reading frame protein. Human secreted protein clone yc25_1 protein sequence SEQ ID NO:26. KCNA3.1B Human ALP.	899  460  188  173  148  430  283  157  181  439  497  605  647  300  137  688	97 86 74 94 56 61 96 43 73 91 84 76
272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738  AK024397  AF141326  AF156530  Y29336  Y73402  AF016411	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Arabidopsis thaliana Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus Homo sapiens Homo sapiens	VPS10 domain receptor protein SORCS  ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein Contains PF]00069 Eukaryotic protein kinase domain. protein-tyrosine phosphatase unnamed protein product RNA helicase HDB/DICE1 ETS-domain transcriptional repressor PE1 Human secreted protein clone cs756_2 alternate reading frame protein. Human secreted protein clone yc25_1 protein sequence SEQ ID NO:26. KCNA3.1B Human ALP. differentiation enhancing factor 1	899 460 188 173 148 430 283 157 181 439 497 605 647 300	97 86 74 94 56 61 96 43 73 91 84 76 100
_,_	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738  AK024397  AF141326  AF156530  Y29336  Y73402  AF016411  W89253  AF112886	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Bos taurus	VPS10 domain receptor protein SORCS  ionotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein Contains PF]00069 Eukaryotic protein kinase domain. protein-tyrosine phosphatase unnamed protein product RNA helicase HDB/DICE1 ETS-domain transcriptional repressor PE1 Human secreted protein clone cs756_2 alternate reading frame protein. Human secreted protein clone yc25_1 protein sequence SEQ ID NO:26. KCNA3.1B Human ALP. differentiation enhancing factor 1	899  460  188  173  148  430  283  157  181  439  497  605  647  300  137  688  750	97 86 74 94 56 61 96 43 73 91 84 76 100 90
272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288	AF137367  L34938  AL022724  AF265555  G02872  L40380  AB046851  AC008075  M83738  AK024397  AF141326  AF156530  Y29336  Y73402  AF016411  W89253	tabacum Mus musculus Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	VPS10 domain receptor protein SORCS lonotropic glutamate receptor  dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1) ubiquitin-conjugating BIR-domain enzyme APOLLON Human secreted protein, SEQ ID NO: 6953. thyroid receptor interactor KIAA1631 protein Contains PF100069 Eukaryotic protein kinase domain. protein-tyrosine phosphatase unnamed protein product RNA helicase HDB/DICE1 ETS-domain transcriptional repressor PE1 Human secreted protein clone cs756_2 alternate reading frame protein. Human secreted protein clone yc25_1 protein sequence SEQ ID NO:26. KCNA3.1B Human ALP.	899  460  188  173  148  430  283  157  181  439  497  605  647  300  137  688	97 86 74 94 56 61 96 43 73 91 84 76 100 90

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman	% Identity
NO.	<del></del>	norvegicus	<del> </del>	Score	<del></del>
292	AF102854	Rattus		+104	<del> </del>
292	AF102834	norvegicus	membrane-associated guanylate kinase-	124	53
293	X99211	Drosophila melanogaster	interacting protein 2 Maguin-2 ubiquitin-specific protease	143	38
294	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	185	94
295	Y94890	Homo sapiens	Human protein clone HP02798.	108	59
296	AF019767	Homo saplens	zinc finger protein	154	96
297	Y28568	Homo sapiens	Secreted peptide clone bd577 1.	568	84
298	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	182	97
299	B08906	Homo sapiens	Human secreted protein sequence encoded by gene 16 SEQ ID NO:63.	605	69
300	R58890	Homo sapiens	Human-32 cadherin-related molecule.	212	97
301	AF022859	Homo sapiens	neuropilin-2(a0)	277	100
302	Y71124	Homo sapiens	Human mitogenic regulator duox2.	716	97
303	Y44297	Homo sapiens	Human receptor tyrosine kinase.	228	97
304	D32050	Homo sapiens	alanyl-tRNA synthetase	192	80
305	U43586	Homo sapiens	protein kinase related to Raf protein kinases; Method: conceptual translation supplied by author	428	72
306	R54872	Homo sapiens	Human H13 viral receptor mutant 4.	280	95
307	D78572	Mus musculus	membrane glycoprotein	199	41
308	AF255614	Rattus norvegicus	scaffolding protein SLIPR	639	88
309	S79463	Mus sp.	semaphorin homolog-M-Sema F	162	89
310	AF178941	Homo sapiens	ATP-binding cassette sub-family A member 2	736	100
311	U03413	Dictyostelium discoideum	calcium binding protein	151	36
312	Y87347	Homo sapiens	Human signal peptide containing protein HSPP- 124 SEQ ID NO:124.	744	100
313	Z97055	Homo sapiens	dJ388M5.4 (putative GS2 like protein)	789 .	99
314	AC004010	Homo sapiens	similar to Leucine-rich transmembrane proteins; 44% similarity to U42767 (PID:g1736918)	197	38
315	AL021392	Homo sapiens	dJ439F8.2 (supported by GENSCAN and GENEWISE)	278	38
316	U70209	Mus musculus	polycystic kidney disease I protein	165	38
317	AF109643	Rattus norvegicus	coxsackie-adenovirus-receptor homolog	223	38
318 319	AF104923	Homo sapiens	putative transcription factor	138	84
	AF100287	Trypanosoma vivax	activated protein kinase C receptor homolog	141	38
320	G00588	Homo sapiens	Human secreted protein, SEQ ID NO: 4669.	125	51
321 322	Y21591 D26070	Homo sapiens Homo sapiens	Human secreted protein (clone CC332-33). human type 1 inositol 1,4,5-trisphosphate receptor	459 232	97 97
323	Y27918	Homo sapiens	Human secreted protein encoded by gene No. 123.	306	88
324	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	209	70
325	M19650	Homo sapiens	2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37)	214	97
326	W80396	Homo sapiens	A secreted protein encoded by clone bp646_10.	140	70
327	X75756	Homo sapiens	protein kinase C mu	540	78
328	G02292	Homo sapiens	Human secreted protein, SEQ ID NO: 6373.	721	99
329	AF168990	Homo sapiens	putative GTP-binding protein	877	99
330	S67984	Homo sapiens	anti-HIV gpl 20 antibody heavy chain variable region	581	80
331	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	2823	98
332	Y87330	Homo sapiens	Human signal peptide containing protein HSPP- 107 SEQ ID NO:107.	1127	100
333	Y28503	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	320	98
334	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95%	327	93

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
			similarity to P49205 (PID:g1345860)		
335	Y87347	Homo sapiens	Human signal peptide containing protein HSPP- 124 SEQ ID NO:124.	1111	67
336	AF006466	Mus musculus	lymphocyte specific formin related protein	193	75
337	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	632	97
338	Y13443	Homo sapiens	Amino acid sequence of hSlo3-2.	516	100
339	Y07637	Homo sapiens	putative GABA-gated chloride channel	189	100
340	Y05734	Homo sapiens	Human Grb7 effector 2.2412 protein.	2156	99
341	AE000497	Escherichia coli	L-idonate transcriptional regulator	928	98
342	D90855	Escherichia coli	glycerol-3-phosphate dehydrogenase (EC 1.1.99.5) chain A, anaerobic	769	99
343	D85613	Escherichia coli	membrane component	399	100
344	M93239	Escherichia coli	transmembrane protein	232	100
345	M60177	Escherichia coli	enterobactin	759	99
346	D90699	Escherichia coli	Sensor protein copS (EC 2.7.3).	638	97
347	D90843	Escherichia coli	CapB protein.	552	100
348	M13422	Escherichia	49 kd protein	1193	96
349	L10328	coll   Escherichia	similar to drug resistance translocases	340	90
350	X69942	Mus .	enhancer-trap-locus-1	560	82
351	AF239613	Homo sapiens	apamin-sensitive small-conductance Ca2+-	463	80
			activated potassium channel	L	<u> </u>
352	D90777	Escherichia coli	3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157) (b- hydroxybutyryl-CoA dehydrogenase) (BhbD).	577	100
353	D90863	Escherichia coli	similar to	311	98
354	Y52386	Homo sapiens	Human transmembrane protein HP02000.	133	58
355	Y31645	Homo sapiens	Human transport-associated protein-7 (TRANP-7).	482	55
356	Y58637	Homo sapiens	Protein regulating gene expression PRGE-30.	119	51
357	AF119226	Homo sapiens	dual-specificity tyrosine phosphatase YVH1	1788	100
358	Y87219	Homo sapiens	Human secreted protein sequence SEQ ID NO:258.	165	-100
359	J00132	Homo sapiens	beta-fibrinogen	233	93
360	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	128	70
361	R28916	Homo sapiens	Type III procollagen (prior art).	108	40
362	U16655	Rattus norvegicus	phospholipase C delta-4	649	65
363	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	95	42
364	U47276	Gallus gallus	chicken brain factor-2	104	34
365	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	183	65
366	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	118	46
367	X98258	Homo sapiens	M-phase phosphoprotein 9	564	75
368	AL021366	Homo sapiens	c1CK0721Q.3 (Kinesin related protein)	3387	99
369	U70932	Peromyscus leucopus	reverse transcriptase	92	59
370	X86400	Homo sapiens	gamma subunit of sodium potassium ATPase like	242	73
371	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	165	56
372	U49974	Homo sapiens	mariner transposase	257	55
373	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	21193	99
374	AF234765	Rattus norvegicus	serine-arginine-rich splicing regulatory protein SRRP86	1182	78
375	U49974	Homo sapiens	mariner transposase	172	55

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
376	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	221	67
377	G00669	Homo sapiens	Human secreted protein, SEQ ID NO: 4750.	600	100
378	X52574	Mus musculus	GTP binding protein	1456	91
379	R69095	Homo sapiens	Anti-HIV Fab tat31 light chain.	68	37
380	J04974	Homo sapiens	alpha-2 type XI collagen	125	37
381	AB002405	Homo sapiens	LAK-4p	530	43
382	U64830	Dictyostelium discoideum	protein tyrosine kinase	115	44
383	G02916	Homo sapiens	Human secreted protein, SEQ ID NO: 6997.	618	98
384	G01194	Homo sapiens	Human secreted protein, SEQ ID NO: 5275.	617	93
385	AJ245822	Homo sapiens	type I transmembrane receptor	4560	100
386	D86974	Homo sapiens	KIAA0220	2148	98
387	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	142	50
388	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	99	59
389	M12140	Homo sapiens	envelope protein 9	197	51
390	AJ293309	Homo sapiens	NHP2 protein	461	77
391	Y42751	Homo sapiens	Human calcium binding protein 2 (CaBP-2).	181	94
392	W48351	Homo saplens	Human breast cancer related protein BCRB2.	241	66
393	Y14442	Homo sapiens	olfactory receptor protein	339	54
394	W85607	Homo sapiens	Secreted protein clone da228 6.	957	100
395	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	171	34
396	G03930	Homo sapiens	Human secreted protein, SEQ ID NO: 8011.	250	100
397	AB032904	Hylobates syndactylus	dopamine receptor D4	105	35
398	AJ007798	Homo sapiens	stromal antigen 3, (STAG3)	861	.85
399	¥91405	Homo sapiens	Human secreted protein sequence encoded by gene 2 SEQ ID NO:126.	1047	92
400	Y29861	Homo sapiens	Human secreted protein clone cb98 4.	162	37
401	D87002	Homo sapiens	similar to rat integral membrane glycoprotein; accession number 221513.	527	78
402	AF100754	Homo sapiens	ancient ubiquitous protein AUP1 isoform	853	95
403	X74904	Gallus gallus	alpha-2-macroglobulin receptor	258	60
404	AF075462	Mus musculus	ADP-ribosylation factor-directed GTPase activating protein isoform b	545	89
405	X92887	Human endogenous retrovirus K	pol/env	162	30
406	Y30162	Homo sapiens	Human dorsal root receptor 4 hDRR4.	325	72
407	AK.022626	Homo sapiens	unnamed protein product	2833	99
408	L13802	Homo sapiens	ribosmal protein small subunit	264	92
409	Y91600	Homo sapiens	Human secreted protein sequence encoded by gene 9 SEQ ID NO:273.	1788	89
410	W88745	Homo sapiens	Secreted protein encoded by gene 30 clone HTSEV09.	2004	99
411	AB043953	Mus musculus	Chat-H	2628	82
412	Y86233	Homo sapiens	Human secreted protein HNTMX29, SEQ ID NO:148.	1014	92
413	U10542	Pan troglodytes	MHC class I A	265	71
414	AF155097	Homo sapiens	NY-REN-7 antigen	850	95
415	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	88	48
416	Y57911	Homo sapiens	Human transmembrane protein HTMPN-35.	266	89
417	W27651	Homo sapiens	Secreted protein AT205.	481	60
418	Y76884	Homo sapiens	Retinoblastoma binding protein-7 sequence.	3077	87
419	AF255559	Notothenia coriiceps	alpha tubulin	289	68
420	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	209	74
421	AL109827	Homo sapiens	dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4)))	1446	96
422	AC008075	Arabidopsis thaliana	F24J5.4	112	35

SEO	Accession	Species	Description	Smith-	1%
ID ID	No.	Dptotts	Description	Waterman	Identity
NO:	1.0.	1		Score	Identity
423	AF231705	Homo sapiens	Alu co-repressor 1	1090	100
424	AF234887	Homo sapiens	FLAMINGO 1		
425	Y35942	Homo sapiens	1	6268	97
723	133742	nonto sapicais	Extended human secreted protein sequence, SEQ ID NO. 191.	1961	99
426	AB009288	Homo sapiens	N-copine	635	<del> </del>
427	L12392	Homo sapiens	Huntington's Disease protein		98
428	Y94990	Homo sapiens		16080	99
429	AJ293573	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	768	98
430	Y84441		zinc finger protein Cezanne	542	87
430	104441	Homo sapiens	Amino acid sequence of a human RNA-	2074	100
431	G02850	Homo sapiens	associated protein.		
432	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 6931.	723	95
432	AF159296		Human secreted protein, SEQ ID NO: 8148.	73	42
433	AF 139290	Lycopersicon esculentum	extensin-like protein	613	48
434	W48351		W	100	
434	X73874	Homo sapiens	Human breast cancer related protein BCRB2.	135	44
435	AF161426	Homo sapiens	phosphorylase kinase	3442	97
437	Y30812	Homo sapiens	HSPC308	268	74
437	G03798	Homo sapiens	Human secreted protein encoded from gene 2.	1055	52
438	X14766	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	168	56
		Homo sapiens	GABA-A receptor alpha 1 subunit	2294	96
440 441	X02344 AF168418	Homo sapiens	beta-tubulin	311	95
		Homo sapiens	activating signal cointegrator 1	1882	100
442	L11672 G03203	Homo sapiens	zinc finger protein	795	54
444	A52140	Homo sapiens unidentified	Human secreted protein, SEQ ID NO: 7284.	93	26
444	X98330		HUMAN NDR	2451	100
446	AF1 16712	Homo sapiens	ryanodine receptor 2	9356	99
440		Homo sapiens	PRO2738	227	49
447	AF245447 AF133086	Homo sapiens	sphingosine kinase type 2 isoform	576	99
449	U87305	Homo sapiens	membrane-type serine protease 1	2630	94
449	08/305	Rattus norvegicus	transmembrane receptor UNC5H1	817	93
450	AF081249	Homo sapiens	TANZ -load - wai MONTEA La inc	4550	-
451	AC005498	Homo sapiens	JAW1-related protein MRVIIA long isoform R31665 1	4568 316	99
452	M60235	Homo sapiens	granule membrane protein-140	464	62
453	AB036706	Homo sapiens	intelectin	730	73 88
454	G00918	Homo sapiens	Human secreted protein, SEQ ID NO: 4999.	263	81
455	Y22634	Homo sapiens	Human cytokine inducible regulatory protein-1	192	67
433	122004	House sapicus	(CIRP-1).	192	0/
456	Y36705	Homo sapiens	Fragment of human secreted protein encoded by	106	40
	1 200.05	Troine supreus	gene 62.	100	1 40
457	N91325	Homo sapiens	DNA encoding human growth hormone receptor.	3282	96
458	M19155	Plasmodium	S-antigen precursor	110	36
		falciparum		1.0	1 30
459	Y13377	Homo sapiens	Amino acid sequence of protein PRO257.	509	98
460	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	149	43
			clone HTDAD22.	177	~
461	Y14482	Homo sapiens	Fragment of human secreted protein encoded by	184	54
	1		gene 17.	***	~
462	Y53005	Homo sapiens	Human secreted protein clone pm749 8 protein	135	47
	**		sequence SEQ ID NO:16.		"
463	X84960	Triticum	low molecular weight glutenin	109	33
		aestivum	· · · · · · · · · · · · · · · · · · ·	.07	33
464	W19919	Homo sapiens	Human Ksr-1 (kinase suppressor of Ras).	1781	-85
465	AF189764	Mus	alpha/beta hydrolase-1	502	59
		musculus		202	37
466	U93569	Homo sapiens	р40	101	30
467	Y41528	Homo sapiens	Fragment of human secreted protein encoded by	1172	99
		and orbiding	gene 77.	11/4	رر
468	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	149	52
	AJ000008	Homo sapiens	PI3-kinase	5832	97
469	,				47
469	X70922	i Mne !			
469	X70922	Mus	neurotoxin homologue	118	<b>*</b> /
	X70922 G03797	Mus musculus Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	198	75

SEQ	Accession	Species	Description	Smith-	1%
ID	No.			Waterman	Identity
NO:			<u> </u>	Score	'
		1	gene 62.	<u> </u>	
473	G02313	Homo sapiens	Human secreted protein, SEQ ID NO: 6394.	328	100
474	Y07007	Homo sapiens	Breast cancer associated antigen precursor	1013	97
		Ĭ	sequence.	1	-
475	W93254	Homo sapiens	Human ESRP1 protein.	943	80
476	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	236	65
477	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	202	60
	1		clone HTDAD22.	1 ~02	00
478	G01870	Homo sapiens	Human secreted protein, SEQ ID NO: 5951.	267	100
479	AF102777	Mus	FYVE finger-containing phosphoinositide kinase	3427	92
		musculus	1 1 1 1 Imger-containing phospholicalitie kinase	3427	1 32
480	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	123	53
481	W87701	Homo sapiens	A human membrane fusion protein designated	221	
102	7,07,01	Momo sapicus	SYTAX1.	221	77
482	G03119	Homo sapiens		101	100
483	AF210651	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	131	39
484	AF010144		NAG18	124	59
<del>484</del> 485	G00637	Homo sapiens	neuronal thread protein AD7c-NTP	343	50
		Homo sapiens	Human secreted protein, SEQ ID NO: 4718.	129	70
486	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	149	73
407	V96169	<del>                                      </del>	3		ļ
487	Y76167	Homo sapiens	Human secreted protein encoded by gene 44.	627	100
488	AJ275213	Homo sapiens	stabilin-1	1244	91
489	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	313	65
490	L12392	Homo sapiens	Huntington's Disease protein	16081	100
491	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	197	66
492	J03799	Homo sapiens	laminin-binding protein	228	70
493	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	128	41
494	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	197	67
495	AC005175	Homo sapiens	R31449 3	889	94
496	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	229	61
497	AB030237	Canis	D4 dopamine receptor	90	48
		familiaris	27 dopamino recopior	70	40
498	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	228	65
499	U70935	Peromyscus	reverse transcriptase	213	52
	0.0322	maniculatus	10 voise danser puse		32
500	U48508	Homo sapiens	skeletal muscle ryanodine receptor	26406	99
501	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	105	58
502	AF119851	Homo sapiens	PRO1722	156	62
503	AF113685	Homo sapiens	PRO0974		
504	U79458	Homo sapiens	WW domain binding protein-2	116 322	50
505	W29651	Homo sapiens	Human secreted protein CD124 3.		59
506	W85459	Homo sapiens	Secreted protein encoded by clone dh1135 9.	608	55
507	Y86265		William and and and and and and and and and and	986	70
		Homo sapiens	Human secreted protein HUSXE77, SEQ ID NO:180.	115	33
508	AL160175	Homo sapiens	bA243J16.3 (similar to MYLK (myosin, light polypeptide kinase))	184	92
509	U43360	Peromyscus maniculatus	reverse transcriptase	97	62
510	G03789		W		
11		Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	117	63
	W79092	Homo sapiens	Human secreted protein dn740_3.	1058	100
512	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	205	64
13	AJ133439	Homo sapiens	GRIP1 protein	2151	100
514	AE003456	Drosophila melanogaster	CG6393 gene product	259	42 .
515	Z17206	Xenopus laevis	p46XIEg22	128	40
16	AF104413	Homo sapiens	large tumor suppressor 1	1266	04
517	G03797	Homo sapiens		1766	94
	AF151083		Human secreted protein, SEQ ID NO: 7878.	92	40
110	1 WLT31093	Homo sapiens	HSPC249	444	98
	C90064	Uama			
518 519 520	S80864 X92485	Homo sapiens Plasmodium	cytochrome c-like polypeptide pval	318 170	50 61

SEQ	Accession	Species	Description	Smith-	1%
D `	No.	1	1	Waterman	Identity
NO:	1	1	·	Score	reculary
521	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	159	59
522	AF121857	Homo sapiens	sorting nexin 7	259	40
523	G02654	Homo sapiens	Human secreted protein, SEQ ID NO: 6735.	82	37
524	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	253	73
, 52.	1,0002,	1	HPMBQ32.	233	/3
525	AF119851	Homo sapiens	PRO1722	162	57
526	Y27761	Homo sapiens	Human secreted protein encoded by gene No. 47.	154	57
527	G02707	Homo sapiens	Human secreted protein, SEQ ID NO: 6788.	70	45
528	U47924	Homo sapiens	C8	1112	86
529	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	84	45
530	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	111	60
531	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	92	65
532	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	75	29
533	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	182	48
534	AF068286	Homo sapiens	HDCMD38P	861	100
535	U07707	Homo sapiens	epidermal growth factor receptor substrate	228	60
536	G01955	Homo sapiens	Human secreted protein, SEQ ID NO: 6036.	484	75
537	AF219232	Gallus gallus	qin-induced kinase	206	53
538	AF135022	Homo sapiens	mediator	128	100
539	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	141	59
540	AF016430	Caenorhabditi s elegans	contains similarity to a BR-C/TTK domain	853	39
541	AC003093	Homo sapiens	OXYSTEROL-BINDING PROTEIN; 45% similarity to P22059 (PID:g129308)	408	66
542	M29487	Homo sapiens	integrin alpha subunit precursor	517	81
543	AF102530	Mus musculus	olfactory receptor F3	327	73
544	Y73431	Homo sapiens	Human secreted protein clone yb186_1 protein sequence SEQ ID NO:84.	386	100
545	AE004833	Pseudomonas aeruginosa	probable TonB-dependent receptor	279	42
546	G03793	Homo sapiens	Human secreted protein, SEQ ID NO: 7874.	264	53
547	Y69192	Homo sapiens	A human monocyte-macrophage apolipoprotein B receptor protein.	1772	67
548 ·	Y91493	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:166.	176	100
549	G01571	Homo sapiens.	Human secreted protein, SEQ ID NO: 5652.	777	99
550	AF044588	Homo sapiens	protein regulating cytokinesis 1; PRC1	1953	88
551	Y29332	Homo sapiens	Human secreted protein clone pe584_2 protein	1224	94
550	700220	<del>                                     </del>	sequence.		
552 553	X98330 Y42782	Homo sapiens	ryanodine receptor 2	24621	99
554	AB025258	Homo sapiens Mus	Human UC Band #331 protein.	684	95
JJ4	ADV23238	Mus musculus	granuphilin-a	501	41
555	AJ010346	Homo sapiens	RING-H2	1468	100
556	W92388	Homo sapiens	Human TR-interacting protein \$239a.	538	92
557	AF119851	Homo sapiens	PRO1722	175	59
558	AF117756	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP150	183	32
559	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	319	68
560	D86214	Mus musculus	Ca2+ dependent activator protein for secretion	1010	93
561	AF187325	Canis familiaris	meianoma antigen	287	55
562	AJ001981	Homo sapiens	OXAIL	2512	00
563	Z17238	Rattus	glutamate receptor subtype delta-1	2512 338	99 66
564	W30638	norvegicus Homo sapiens	Partial human 7-transmembrane receptor	371	100
565	AC005000		HAPO167 protein.		
565	AC005620	Homo sapiens	R33590_1	467	97
566	Y99358	Homo sapiens	Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63.	1138	78
567	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	1002	58
568	AF151043	Homo sapiens	HSPC209	798	100

SEO	Accession	Species	Description	Smith-	T%
ID `	No.			Waterman	Identity
NO:	1		<b>!</b>	Score	100min
569	AF097518	Homo sapiens	liver-specific transporter	231	100
570	AB035698	Homo sapiens	Misshapen/NIK-related kinase MINK-1	1532	100
571	Y07096	Homo sapiens	Colon cancer associated antigen precursor	1064	100
371	1 20,050	nomo sapiciis	sequence.	1004	100
572	AL031177	Homo sapiens	dJ889M15.3 (novel protein)		\ <u></u>
573	Y66639	Homo sapiens		735	55
574			Membrane-bound protein PRO290.	254	45
575	AB037108	Homo sapiens	seven transmembrane domain orphan receptor	1883	99
	D43949	Homo sapiens	This gene is novel.	836	100
576	Y48596	Homo sapiens	Human breast tumour-associated protein 57.	108	50
577	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	141	75
578	R95913	Homo sapiens	Neural thread protein.	140	65
579	AK025116	Homo sapiens	unnamed protein product	201	70
580	Y86473	Homo sapiens	Human gene 52-encoded protein fragment, SEQ	77	70
		<u> </u>	ID NO:388.	1	1
581	AF196779	Homo sapiens	JM10 protein	450	100
582	AF188706	Homo sapiens	g20 protein	330	98
583	AB030234	Canis	D4 dopamine receptor	64	56
	1	familiaris		1	1
584	G02621	Homo sapiens	Human secreted protein, SEQ ID NO: 6702.	345	90
585	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	268	85
	1		Antigen)	1	55
586	Y30819	Homo sapiens	Human secreted protein encoded from gene 9.	235	35
587	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	132	56
588	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	182	79
589	AF235017	Mus	2P1 protein	764	80
507	111 255017	musculus	21 1 protein	1 /64	∾
590	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	329	
370	1100027	Homo Sapiens	HPMBQ32.	329	81
591	Y30709	Homo sapiens	Amino acid sequence of a human secreted	110	42
331	130709	moino sapiens	protein.	110	43
592	Y53875	Home conices	L*	1000	
392	1330/3	Homo sapiens	A human seven transmembrane signal transducer	1369	92
593	Y53051	17	polypeptide.		ļ
293	103031	Homo sapiens	Human secreted protein clone dd119_4 protein	1112	97
594	Y27658	ļ.,	sequence SEQ ID NO:108.		
595	G03798	Homo sapiens	Human secreted protein encoded by gene No. 92.	763	79
		Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	156	58
596	AF151110	Mus	COP1 protein	2215	95
£08	600,506	musculus			
597	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	157	65
598 <sup>.</sup>	AF192499	Mus	putative secreted protein ZSIG37	143	40
	· · · · · · · · · · · · · · · · · · ·	musculus			
599	AF119855	Homo sapiens	PRO1847	236	76
600.	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	212	73
601	Y00295	Homo sapiens	Human secreted protein encoded by gene 38.	567	88
602	AF184971	Homo sapiens	class II cytokine receptor ZCYTOR7	2015	74
603	AF061936	Homo sapiens	diacylglycerol kinase iota	773	96
604	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	1333	93
			Antigen)		
605	AB033106	Homo sapiens	KIAA1280 protein	3915	100
606	X75756 ·	Homo sapiens	protein kinase C mu	3916	99
607	D86983	Homo sapiens	similar to D.melanogaster peroxidasin(U11052)	5758	99
608	W69341	Homo sapiens	Secreted protein of clone CG279_1.	1377	99
609	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	339	82
	""""	Tomo achiona	HPMBQ32.	227	02
610	Y27868	Homo sapiens	Human secreted protein encoded by gene No.	116	62
3.0	12,000	Lionio sapiens	107.	116	02
611	AF202636	Homo sapiens		23.64	100
612	AF090944		angiopoletin-like protein PP1158	2164	100
		Homo sapiens	PRO0663	218	82
613	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	195	59
614	1 1000000	<del>   </del>	clone HTDAD22.		
614	M87053	Rattus	lens membrane protein	450	84
712	1	norvegicus			
615 616	AC004232	Homo sapiens	FPM315	163	37
	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	205	79

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
617	Y91524	Homo sapiens	Human secreted protein sequence encoded by gene 74 SEQ ID NO:197.	821	99
618	AJ245621	Homo sapiens	CTL2 protein	2258	99
619	Y76198	Homo sapiens	Human secreted protein encoded by gene 75.	108	64
620	AF067864	Homo sapiens	transferrin receptor 2 alpha	3922	94
621	D90721	Escherichia coli	Transmembrane protein dppC	573	90
622	W75858	Homo sapiens	Human secretory protein of clone CS752-3.	730	100
623	Y94982	Homo sapiens	Human secreted protein vb12_1, SEQ ID NO:4.	733	100
624	AF034745	Mus musculus	LNXp80	637	83
625	U42580	Paramecium bursaria Chlorella virus 1	Pro-rich, IPPPNMSLPLS (3x)	94	46
626	U79260	Homo sapiens	unknown	194	70
627	R95913	Homo sapiens	Neural thread protein.	99	50
628	G03450	Homo sapiens	Human secreted protein, SEQ ID NO: 7531.	427	100
629	Y36281	Homo sapiens	Human secreted protein encoded by gene 58.	590	100
630	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	165	76
631	G02139	Homo sapiens	Human secreted protein, SEQ ID NO: 6220.	268	96
632	U16996	Homo sapiens	protein tyrosine posphatase	351	80
633	AF121857	Homo sapiens	sorting nexin 7	2019	100
634	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10 encoded by GenBank Accession Number L25899	340	77
635	¥07090	Homo sapiens	Renal cancer associated antigen precursor sequence.	277	64
636	AB013382	Homo sapiens	DUSP6	414	76
637	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	315	71
638	M95762	Rattus norvegicus	GABA transporter	924	89
639	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	219	60
640	Y01400	Homo sapiens	Secreted protein encoded by gene 18 clone HNHF029.	137	79
641	AC008075	Arabidopsis thaliana	F24J5.4	121	33
642	W74824	Homo sapiens	Human secreted protein encoded by gene 96 clone HAQBK61.	615	62
643	AB015982	Homo sapiens	serine/threonine kinase	485	98
644	Y25806	Homo sapiens	Human secreted protein fragment encoded from gene 23.	162	46
645	AF122904	Homo sapiens	membrane protein DAP10	474	100
646	AF233323	Homo sapiens	Fas-associated phosphatase-1	200	38
647	W48804	Homo sapiens	Homo sapiens clone BK158_1 protein.	1203	99
648	AF257330	Homo sapiens	COBW-like protein	1440	98
649	Y36203	Homo sapiens	Human secreted protein #75.	233	73
650 651	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	173	78
	Y32199	Homo sapiens	Human receptor molecule (REC) encoded by Incyte clone 2022379.	1012	100
652	AB032909	Hylobates agilis	dopamine receptor D4	122	32
722	4 77 60	<del> </del>			
653	AK021848	Homo sapiens	unnamed protein product	186	69
654	W73411	Homo sapiens	Human secreted protein encoded by Gene No. 15.	57	37
654 655			Human secreted protein encoded by Gene No.		
654 655 656	W73411 L22455 G03112	Homo sapiens Rattus	Human secreted protein encoded by Gene No. 15.	57	37
654 655	W73411 L22455 G03112 G02345	Homo sapiens Rattus norvegicus	Human secreted protein encoded by Gene No. 15. mu opioid receptor	116	37
654 655 656 657 658	W73411 L22455 G03112 G02345 W88627	Homo sapiens Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens	Human secreted protein encoded by Gene No. 15. mu opioid receptor  Human secreted protein, SEQ ID NO: 7193. Human secreted protein, SEQ ID NO: 6426. Secreted protein encoded by gene 94 clone HPMBQ32.	116 110 459 291	37 34 45 97 75
654 655 656 657	W73411 L22455 G03112 G02345	Homo sapiens  Rattus norvegicus Homo sapiens Homo sapiens	Human secreted protein encoded by Gene No. 15. mu opioid receptor  Human secreted protein, SEQ ID NO: 7193. Human secreted protein, SEQ ID NO: 6426. Secreted protein encoded by gene 94 clone	116 110 459	37 34 45 97

SEQ	Accession	Species	Description	Smith-	%
ID .	No.	Į.		Waterman	Identity
NO:	l	<u> </u>		Score	<u> </u>
661	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	168	68 .
662	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	375	43
	<u> </u>	<u> </u>	designated HSCOP-6.		
663	W75771	Homo sapiens	Human GTP binding protein APD08.	629	100
664	AL096770	Homo sapiens	bA150A6.2 (novel 7 transmembrane receptor	480	55
			(rhodopsin family) (olfactory receptor like)		
115	A TO 0 0 0 0 0 0 4		protein (hs6M1-21))	070	100
665	AB037734	Homo sapiens	KIAA1313 protein	978	96
666	W82841	Homo sapiens	Human cerebral protein-1.	192	84
667	W82841	Homo sapiens	Human cerebral protein-1.	182	87
668	AB030184	Mus musculus	contains transmembrane (TM) region and ATP binding region	757	68
669	AB032919	Hylobates	dopamine receptor D4	85	37
005	AD032919	muelleri	dopanine receptor 124	1 02	3'
670	AF107295	Rattus	outer membrane protein	746	81
070	AL 10/255	norvegicus	outer memorane process	1,70	} "
671	Z33642	Homo sapiens	leukocyte surface protein	394	93
672	W85608	Homo sapiens	Secreted protein clone du410 5.	261	91
673	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	106	48
674	AL035587	Homo sapiens	dJ475N16.4 (KJAA0240)	2388	99
675	Y59668	Homo saplens	Secreted protein 108-005-5-0-C1-FL.	1134	53
676	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	174	74
677	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory	1013	95
			subunit precursor, PDPr		
678	L11625	Mus	receptor protein-tyrosine kinase	545	96
		musculus	,	ł	
679	AL031427	Homo sapiens	dJ167A19.3 (novel protein)	745	100
680	AJ133430	Mus	olfactory receptor	528	77
	<u> </u>	musculus			
681	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	179	70
682	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	336	76
683	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein	118	100
		<u> </u>	sequence SEQ ID NO:92.		<u> </u>
684	U43360	Peromyscus	reverse transcriptase	100	37
202	500005	maniculatus	11		100
685 686	G00885 AK001518	Homo sapiens	Human secreted protein, SEQ ID NO: 4966.	162 590	100
687	G01982	Homo sapiens Homo sapiens	unnamed protein product Human secreted protein, SEQ ID NO: 6063.	718	100
688	Y92241		Human cancer associated antigen precursor	2405	99
080	192241	Homo sapiens	· (MO-REN-46).	2405	עכ
689	AC024792	Caenorhabditi	contains similarity to TR:P78316	423	36
003	110021172	s elegans	Commission of Transport	''	1 **
690	Y27868	Homo sapiens	Human secreted protein encoded by gene No.	183	81
	}		107.		
691	Y56514	Homo sapiens	Human Jurkat cell clone P2-15 AIM10 longest	180	88
	1	•	ORF protein sequence.		1 .
692	¥27795	Homo sapiens	Human secreted protein encoded by gene No. 79.	1539	99
			Human secreted protein encoded by gene 45.	428	-98
693	Y36268	Homo sapiens	i Fluman secreted protein encoded by gene 45.		
	Y36268 U12465	Homo sapiens Homo sapiens	ribosomal protein L35	308	89
				308 1517	99
693 694 695 696	U12465	Homo sapiens	ribosomal protein L35		1
694 695	U12465 Y45272	Homo sapiens Homo sapiens	ribosomal protein L35 Human secreted protein encoded from gene 16. TANK bluding kinase TBK1 Human secreted protein encoded by gene 44	1517	99
694 695 696 697	V12465 Y45272 AF191838 Y02693	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	ribosomal protein L35 Human secreted protein encoded from gene 16. TANK blading kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22.	1517 1242 275	99 98 75
694 695 696 697	U12465 Y45272 AF191838	Homo sapiens Homo sapiens Homo sapiens	ribosomal protein L35 Human secreted protein encoded from gene 16. TANK blading kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP-	1517 1242	99 98
694 695 696 697 698	V12465 Y45272 AF191838 Y02693 Y87280	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	ribosomal protein L35 Human secreted protein encoded from gene 16. TANK blading kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP- 57 SEQ ID NO:57.	1517 1242 275 576	99 98 75
694 695 696 697	V12465 Y45272 AF191838 Y02693	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	ribosomal protein L35 Human secreted protein encoded from gene 16. TANK blading kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP- 57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ	1517 1242 275	99 98 75
694 695 696 697 698	U12465 Y45272 AF191838 Y02693 Y87280 Y97999	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	ribosomal protein L35 Human secreted protein encoded from gene 16. TANK blading kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP- 57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1.	1517 1242 275 576 729	99 98 75 90
694 695 696 697 698 699	U12465 Y45272 AF191838 Y02693 Y87280 Y97999	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	ribosomal protein L35 Human secreted protein encoded from gene 16. TANK blading kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP- 57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1. putative serine/threonine protein kinase	1517 1242 275 576 729	99 98 75 90 99
694 695 696 697 698 699 700 701	U12465 Y45272 AF191838 Y02693 Y87280 Y97999 AJ006701 AF209198	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	ribosomal protein L35 Human secreted protein encoded from gene 16. TANK blinding kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP- 57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1. putative serine/threonine protein kinase zinc finger protein 277	1517 1242 275 576 729 610 2357	99 98 75 90 99 79 100
694 695 696 697 698	U12465 Y45272 AF191838 Y02693 Y87280 Y97999	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	ribosomal protein L35 Human secreted protein encoded from gene 16. TANK blading kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP- 57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1. putative serine/threonine protein kinase	1517 1242 275 576 729	99 98 75 90 99
694 695 696 697 698 699 700 701 702	U12465 Y45272 AF191838 Y02693 Y87280 Y97999 AJ006701 AF209198 AJ298841	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus	ribosomal protein L35 Human secreted protein encoded from gene 16. TANK blinding kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP- 57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1. putative serine/threonine protein kinase zinc finger protein 277 torsinA protein	1517 1242 275 576 729 610 2357 709	99 98 75 90 99 79 100 45
694 695 696 697 698 699 700 701 702	U12465 Y45272 AF191838 Y02693 Y87280 Y97999 AJ006701 AF209198 AJ298841 AK021729	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	ribosomal protein L35 Human secreted protein encoded from gene 16. TANK blinding kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP- 57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1. putative serine/threonine protein kinase zinc finger protein 277 torsinA protein unnamed protein product	1517 1242 275 576 729 610 2357 709	99 98 75 90 99 79 100 45
694 695 696 697 698 699 700 701 702	U12465 Y45272 AF191838 Y02693 Y87280 Y97999 AJ006701 AF209198 AJ298841	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus	ribosomal protein L35 Human secreted protein encoded from gene 16. TANK blinding kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP- 57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1. putative serine/threonine protein kinase zinc finger protein 277 torsinA protein	1517 1242 275 576 729 610 2357 709	99 98 75 90 99 79 100 45

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	G02501		II de la companya de la company	Score	
707	R95326	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 6582.  Tumor necrosis factor receptor 1 death domain	125	58 95
			ligand (clone 2DD).		
708	G03002	Homo sapiens	Human secreted protein, SEQ ID NO: 7083.	125	39
709 710	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	516	98
	M63577	Saccharomyc es cerevisiae	SFP1	131	59
711	AB026291	Rattus norvegicus	acetoacetyl-CoA synthetase	467	85
712	D21211	Homo sapiens	protein tyrosine phosphatase (PTP-BAS, type 3)	368	44
713	AF044033	Marmota marmota	olfactory receptor	615	83
714	G03561	Homo sapiens	Human secreted protein, SEQ ID NO: 7642.	251	100
715	AB033062	Homo sapiens	KIAA1236 protein	1380	100
716	G00577	Homo sapiens	Human secreted protein, SEQ ID NO: 4658.	80	73
717	Y96864	Homo sapiens	SEQ. ID. 37 from WO0034474.	835	99
718	AJ243396	Homo sapiens	voltage-gated sodium channel beta-3 subunit	234	100
719	U47334	Homo sapiens	similar to chicken gamma aminobutyric acid receptor beta4 subunit	578	99
720	AB020598	Homo sapiens	peptide transporter 3	1096	100
721	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	570	74
722	J05046	Homo sapiens	insulin receptor-related receptor	6787	100
723	AF001958	Ambystoma tigrinum	electrogenic Na+ bicarbonate cotransporter; NBC	111	41
724	AF127084	Mus musculus	semaphorin cytoplasmic domain-associated protein 3A	5253	94
725	X54673	Homo sapiens	GABA transporter	3114	99
726	AF016191	Rattus norvegicus	potassium channel	370	100
727	AB029559	Rattus .	BATI	139	35
728	Y28503	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	2186	97
729	AJ011415	Homo sapiens	plexin-B1/SEP receptor	729	56
730	Z93096	Homo sapiens	bK390B3.1 (manic fringe (Drosophila) homolog)	142	68
<i>7</i> 31	Z10062	Homo sapiens	cDNA encoding a human vanilloid receptor homologue Vanilrep1.	675	99
732	AF161382	Homo sapiens	HSPC264	492	94
733	AB029033	Homo sapiens	KIAA1110 protein	3826	99
734	AE000493	Escherichia coli	putative transport protein	592	97
735	AL033379	Homo sapiens	dJ417022.2 (novel 7 transmembrane receptor (rhodopsin family) protein similar to high- affinity lysophosphatidic acid receptor homolog)	2173	99
736	AF132599	Homo sapiens	RANTES factor of late activated T lymphocytes-	245	56
737	X55019	Homo sapiens	acetylcholine receptor delta subunit	883	99.
738	X91906	Homo sapiens	voltage-gated chloride ion channel	1978	100
739	AB026116	Homo sapiens	organic anion transporter 4	1444	98
740	D00570	Mus musculus	open reading frame (196 AA)	83	24
741	W03626	Homo sapiens	Human thyrotropin GPR N-terminal sequence.	118	40
742	U66059	Homo sapiens	V_segment translation product	614	100
743	AF119815	Homo sapiens	G-protein-coupled receptor	2751	99
744 745	X16663	Homo sapiens	haematopoietic lineage cell protein (AA 1-486)	148	93
	W67838	Homo sapiens	Human secreted protein encoded by gene 32 clone HLTCJ63.	448	95
746	W57260	Homo sapiens	Human semaphorin Y.	2414 .	100
747	W21578	Homo sapiens	Alzheimer's disease protein encoded by DNA from plasmid pGCS2232.	968	65
748	Y94935	Homo sapiens	Human secreted protein clone yd218_1 protein sequence SEQ ID NO:76.	622	100
749	AL022238	Homo sapiens	dJ1042K10.5 (novel protein)	314	85
750	G03889	Homo sapiens	Human secreted protein, SEQ ID NO: 7970.	391	87

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
751	AB025258	Mus musculus	granuphilin-a	773	41
752	Y52386	Homo sapiens	Human transmembrane protein HP02000.	900	99
753	Y48586	Homo sapiens	Human breast tumour-associated protein 47.	2527	99
754	AJ272207	Homo sapiens	putative G protein-coupled receptor 92	694	100
755	M85183	Rattus	vasopressin receptor	979	68
		norvegicus	vacopressin receptor	1313	00
756	AF190501	Homo sapiens	leucine-rich repeat-containing G protein-coupled receptor 6	388	71
757	Y02692	Homo sapiens	Human secreted protein encoded by gene 43 clone HTADX17.	461	87
758	Z22535	Homo sapiens	ALK-3	439	98
759	R04932	Homo sapiens	Interferon-gamma receptor segment from clone 39 responsible for binding the target.	564	97
760	W74902	Homo sapiens	Human secreted protein encoded by gene 175 clone HE8BI92.	1217	99
761	G03706	Homo sapiens	Human secreted protein, SEQ ID NO: 7787.	223	88
762	AB020676	Homo sapiens	KIAA0869 protein	4433	99
763	AK026992	Homo sapiens	unnamed protein product	2285	99
764	AF173358	Homo sapiens	glucocorticoid receptor AF-1 coactivator-1	573	100
765	AF268066	Mus	netrin 4	2019	
		musculus	Thomas 4	2019	89
766	Y48585	Homo sapiens	Human breast tumour-associated protein 46.	1169	89
767	AF230378	Mus musculus	interleukin-1 delta	309	45
768	AF121975	Mus musculus	odorant receptor S18	268	62
769	AB008515	Homo sapiens	RanBPM	611	57
770	Y09945	Rattus	putative integral membrane transport protein	458	50
		norvegicus			30
771	AF226731	Homo sapiens	AD026	688	99
772	Y27132	Homo sapiens	Human glioblastoma-derived polypeptide (clone OA004FG).	1384	100
773	X87832	Homo sapiens	NOV/plexin-A1 protein	1821	98
774	AB025258	Mus musculus	granuphilin-a	500	41
775	AF125101	Homo sapiens	HSPC040 protein	232	93
776	G02815	Homo sapiens	Human secreted protein, SEQ ID NO: 6896.	314	95
777	G02493	Homo sapiens	Human secreted protein, SEQ ID NO: 6574.	191	68
778	R03301	Homo sapiens	Sequence of pre-human atrial natriuretic peptide.	213	45
779	AL357374	Homo sapiens	bA353C18.2 (novel protein)	232	100
780	AF100346	Homo sapiens	neuronal voltage gated calcium channel gamma- 3 subunit	1434	89
781	Y19566	Homo sapiens	Amino acid sequence of a human secreted protein.	103	52
782	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	1098	93
783	AF084464	Rattus	GTP-binding protein REM2	141	30
784	W49042	norvegicus	· ·		
		Homo sapiens	Human low density lipoprotein binding protein LBP-3.	2693	99
785	AF238381	Homo sapiens	PTOVI	1904	91
786	Y91870	Homo sapiens	Human apoptosis related protein.	547	100
787	Y71062	Homo sapiens	Human membrane transport protein, MTRP-7.	1062	94
788	AF117754	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	8684	98
789	AL049569	Homo sapiens	dJ37C10.3 (novel ATPase)	2848	96
790	AF151848	Homo sapiens	CGI-90 protein	745	96
791	Y08639	Homo sapiens	nuclear orphan receptor ROR-beta	1421	95
792	Y41706	Homo sapiens	Human PRO381 protein sequence.	644	99
793	AF121228	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP95	1037	100
794	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	124	62
795	Y69384	Homo sapiens	Amino acid sequence of a 14274 receptor	119	100
706	W40015	<del>                                     </del>	protein.		
796	W40215	Homo sapiens	Human macrophage antigen.	1358	99

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	1.		Waterman	Identity
NO:				Score	
797	AF258340	Homo sapiens	hepatocellular carcinoma-associated antigen 112	1151	99
798	AF159615	Homo sapiens	FGF receptor activating protein 1	461	98
799	Y59863	Homo sapiens	Human normal uterus tissue derived protein 26.	797	99
800	W70459	Homo sapiens	Human T1-receptor ligand III splice variant 2.	572	92
801	L00073	Homo sapiens	renin	1913	93
802	P92219	Homo sapiens	CRI protein.	11963	97
803	X15357	(human)	ANTO A	ļ <u> </u>	<u> </u>
804	W64473	Homo sapiens Homo sapiens	ANP-A receptor preprotein (AA -32 to 1029)	5199	98
805	AJ243874	Homo sapiens	Human secreted protein from clone EC172_1. oligophrenin-4	4018	95
806	G01731	Homo sapiens	Human secreted protein, SEQ ID NO: 5812.	2067 284	100
807	Z24680	Homo sapiens	garp	1562	100 83
808	AF171669	Homo sapiens	glycoprotein-associated amino acid transporter	1364	90
	1.4.1.100	Tronic suplois	LAT2	1304	1 30
809	W70321	Homo sapiens	Secreted protein CC198 1.	1154	96
810	W74843	Homo sapiens	Human secreted protein encoded by gene 115	855	99
	•	1	clone HOVBA03.		
811	AF108831	Homo sapiens	K:Cl cotransporter 3	4561	100
812	AF092135	Homo sapiens	PTD014	862	100
813	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10	784	100
			encoded by GenBank Accession Number	1	
			L25899		
814	G01563	Homo sapiens	Human secreted protein, SEQ ID NO: 5644.	330	100
815 816	AF051151	Homo sapiens	Toll/interleukin-1 receptor-like protein 3	3850	99
910	W95630	Homo sapiens	Homo sapiens secreted protein gene clone	358	100
817	G01082	Homo sapiens	gn114_1.  Human secreted protein, SEQ ID NO: 5163.		100
818	AF151800	Homo sapiens	CGI-41 protein	549	100
819	L00352	Homo sapiens	low density lipoprotein receptor	1106 3980	95
820	X04434	Homo sapiens	IGF-I receptor	5832	100
821	G03844	Homo sapiens	Human secreted protein, SEQ ID NO: 7925.	572	100
822	AF212220	Homo sapiens	TERA	396	48
823	Y50125	Homo sapiens	Human glycophosphatidylinositol-anchored	4897	99
•	ĺ	1	protein GPI-122.	1	1
824	AF156778	Homo sapiens	ASB-3 protein	2675	98
825	AF096322	Homo sapiens	neuronal voltage-gated calcium channel gamma-	1105	100
	L		2 subunit		ľ
826	Y07972	Homo sapiens	Human secreted protein fragment #2 encoded	1540	100
007	A D020012	<del>                                     </del>	from gene 28.		
827 828	AB032013 Y13620	Homo sapiens	potassium channel Kv8.1	2435	95
829	Y91474	Homo sapiens	BCL9	5284	96
027	1914/4	Homo sapiens	Human secreted protein sequence encoded by gene 24 SEQ ID NO:147.	541	98
830	X54232	Homo sapiens	glypican	1625	87
831	X14830	Homo sapiens	acetylcholine receptor beta-subunit preprotein	2540	100
832	Y71262	Homo sapiens	Human chondromodulin-like protein, Zchm1.	1002	100
833 ·	G03873	Homo sapiens	Human secreted protein, SEQ ID NO: 7954.	638	96
834	AC003030	Homo sapiens	R29828_1	1389	93
835	Y38422	Homo sapiens	Human secreted protein.	964	87
836	U41557	Caenorhabditi	glycine-rich	85	36
836	U41557	s elegans		<b>8</b> 5	36
		1	dJ1076E17.1 (KIAA0823 protein (continues in	998	<b>36 75</b>
836 837	U41557 AL121889	s elegans Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803))	998	75
836 837 838	U41557  AL121889  AJ011415	s elegans Homo sapiens Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) plexin-B1/SEP receptor	998	75
836 837 838 839	U41557 AL121889 AJ011415 W80398	s elegans Homo sapiens Homo sapiens Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) plexin-B1/SEP receptor A secreted protein encoded by clone cw1543_3.	998 1580 1105	75 60 67
836 837 838 839 840	U41557 AL121889 AJ011415 W80398 G00862	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) plexin-B1/SEP receptor A secreted protein encoded by clone cw1543_3. Human secreted protein, SEQ ID NO: 4943.	998 1580 1105 255	75 60 67 92
836 837 838 839 840 841	U41557 AL121889 AJ011415 W80398 G00862 G02650	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) plexin-B1/SEP receptor A secreted protein encoded by clone cw1543_3. Human secreted protein, SEQ ID NO: 4943. Human secreted protein, SEQ ID NO: 6731.	998 1580 1105 255 644	75 60 67 92 97
836 837 838 839 840 841 842	U41557  AL121889  AJ011415  W80398  G00862  G02650  AF036717	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) plexin-B1/SEP receptor A secreted protein encoded by clone cw1543_3. Human secreted protein, SEQ ID NO: 4943. Human secreted protein, SEQ ID NO: 6731. FGFR signalling adaptor SNT-1	998 1580 1105 255 644 2629	75 60 67 92 97 99
836 837 838 839 840 841	U41557 AL121889 AJ011415 W80398 G00862 G02650	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803))  plexin-B1/SEP receptor  A secreted protein encoded by clone cw1543_3.  Human secreted protein, SEQ ID NO: 4943.  Human secreted protein, SEQ ID NO: 6731.  FGFR signalling adaptor SNT-1  Human secreted protein clone yc27_1 protein	998 1580 1105 255 644	75 60 67 92 97
836 837 838 839 840 841 842	U41557 AL121889 AJ011415 W80398 G00862 G02650 AF036717 Y73446	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803))  plexin-B1/SEP receptor  A secreted protein encoded by clone cw1543_3.  Human secreted protein, SEQ ID NO: 4943.  Human secreted protein, SEQ ID NO: 6731.  FGFR signalling adaptor SNT-1  Human secreted protein clone yc27_1 protein sequence SEQ ID NO:114.	998 1580 1105 255 644 2629 1089	75 60 67 92 97 99 100
836 837 838 839 840 841 842 843	U41557  AL121889  AJ011415  W80398  G00862  G02650  AF036717  Y73446  G02872	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803))  plexin-B1/SEP receptor  A secreted protein encoded by clone cw1543_3.  Human secreted protein, SEQ ID NO: 4943.  Human secreted protein, SEQ ID NO: 6731.  FGFR signalling adaptor SNT-1  Human secreted protein clone yc27_1 protein sequence SEQ ID NO:114.  Human secreted protein, SEQ ID NO: 6953.	998 1580 1105 255 644 2629 1089	75 60 67 92 97 99 100
836 837 838 839 840 841 842	U41557 AL121889 AJ011415 W80398 G00862 G02650 AF036717 Y73446	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803))  plexin-B1/SEP receptor  A secreted protein encoded by clone cw1543_3.  Human secreted protein, SEQ ID NO: 4943.  Human secreted protein, SEQ ID NO: 6731.  FGFR signalling adaptor SNT-1  Human secreted protein clone yc27_1 protein sequence SEQ ID NO:114.	998 1580 1105 255 644 2629 1089	75 60 67 92 97 99 100

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	ļ		Waterman	Identity
NO:	<del> </del>	<del> </del> -	A. A.F.078040 (DVD 2007007)	Score	<del> </del>
848	X99886	Homo sapiens	to AF038969 (PID:g2827207) monocyte chemotactic protein-2	160	76
849	AC005587	Homo sapiens	similar to mouse olfactory receptor 13; similar to	963	98
0.5	11000000	Thomas dapies	P34984 (PID:g464305)	100	100
850	AB038237	Homo sapiens	G protein-coupled receptor C5L2	1767	100
851	AF124490	Homo sapiens	ARF GTPase-activating protein GIT1	3415	98
852	Y86217	Homo sapiens	Human secreted protein HWHGU54, SEQ ID NO:132.	1189	99
853	AF224741	Homo sapiens	chloride channel protein 7	3748	99
854	X17094	Homo sapiens	furin (AA 1-794)	3550	99
855	W78245	Homo sapiens	Fragment of human secreted protein encoded by gene 19.	1245	99
856	R97569	Homo sapiens	Interleukin-2 receptor associated protein p43.	1926	100
857 858	Y41765 AF057306	Homo sapiens	Human PRO1083 protein sequence.	3211	99
859	AK025116	Homo sapiens Homo sapiens	transmembrane proteolipid unnamed protein product	481 374	84 69
860	Y41312	Homo sapiens	Human secreted protein encoded by gene 5 clone	824	100
		1	HLDRM43.		l
862 863	Y25776	Homo sapiens	Human secreted protein encoded from gene 66.	895	99
	Y74188	Homo sapiens	Human prostate tumor EST fragment derived protein #375.	96	30
864	AF167473	Homo sapiens	heme-binding protein	870	99
865 866	G02532 X54870	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 6613.  Type II integral membrane protein	211	67
867	G00700	Homo sapiens	Human secreted protein, SEQ ID NO: 4781.	1201 640	100 99
868	Y07894	Homo sapiens	Human secreted protein fragment encoded from	388	88
869	J00123	Homo sapiens	gene 43. preproenkephalin (	1240	
870	Y91632	Homo sapiens	Human secreted protein sequence encoded by	1349	95 98
		1	gene 25 SEQ ID NO:305.		
871	L04311 Y29988	Homo sapiens	GABA-alpha receptor beta-3 subunit	237	93
872 873	AF161382	Homo sapiens Homo sapiens	Human cytokine family member EF-7 protein. HSPC264	960 1124	94
874	G03412	Homo sapiens	Human secreted protein, SEQ ID NO: 7493.	464	100
875	Y27572	Homo sapiens	Human secreted protein encoded by gene No. 6.	573	96
876	M15530	Homo sapiens	B-cell growth factor	171	56
877	W63681	Homo sapiens	Human secreted protein 1.	1652	99
878	L27867	Rattus norvegicus	neurexophilin	1448	98
879	Y10835	Homo sapiens	Amino acid sequence of a human secreted protein.	321	100
880	W88991	Homo sapiens	Polypeptide fragment encoded by gene 144.	936	100
881	AF118670	Homo sapiens	orphan G protein-coupled receptor	1971	100
882	AF208865	Homo sapiens	EDRF	528	100
883	Y18462	Homo sapiens	cathepsin L	209	72
884	Y94950	Homo sapiens	Human secreted protein clone dh1073_12 protein sequence SEQ ID NO:106.	348	100
885	AF070661	Homo sapiens	HSPC005	404	100
886 887	Y04315 X92744	Homo sapiens Homo sapiens	Human secreted protein encoded by gene 23. hBD-1	385 375	100
888	Y22496	Homo sapiens	Human secreted protein sequence clone cn621 8.	994	94
889	Y41293	Homo sapiens	Human soluble protein ZTMPO-1.	4595 ·	99
890	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	147	63
891	AF208856	Homo sapiens	BM-014	1012	99
892	U29195	Homo sapiens	neuronal pentraxin II	2002	98
893	X68149	Homo sapiens	Burkitt lymphoma receptor 1	1953	100
894	Y94914	Homo sapiens	Human secreted protein clone pw337_6 protein sequence SEQ ID NO:34.	537	100
895	W61630	Homo sapiens	Clone HNFGW06 of EGFR receptor family.	326	63
896	M24110	Homo saplens	G0S19-2 peptide precursor	481	100
897	Z68747	Homo sapiens	imogen 38	2018	99
898	AF186112	Homo sapiens	neurokinin B-like protein ZNEUROK1	619	100
899	AF225420	Homo sapiens	AD025	734	100

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	140.		j.	Score	luctifity
900	P60657	Homo sapiens	Sequence of human lipocortin.	1835	100
901	M27288	Homo sapiens	oncostatin M	1297	99
902	W85737	Homo sapiens	Polypeptide with transmembrane domain.	749	100
903	G01349	Homo sapiens	Human secreted protein, SEQ ID NO: 5430.	650	99
904	Y00261	Homo sapiens	Human secreted protein encoded by gene 4.	1133	99
905	AF039688 AB007836	Homo sapiens	antigen NY-CO-3 Hic-5	771 2544	99 100
906 907	AB007836 AB017507	Homo sapiens Homo sapiens	Apg12	224	100
907	AK000056	Homo sapiens	unnamed protein product	1537	98
909	Y86299	Homo sapiens	Human secreted protein HFOXB55, SEQ ID	427	100
307	100275	Tionio sapions	NO:214.	1 72"	100
910	AF231023	Homo sapiens	protocadherin Flamingo 1	7393	99
911	Y14134	Homo sapiens	Vascular endothelial cell growth inhibitor beta	1319	100
	<u> </u>		protein sequence.	<u></u>	ł
912	Z90420	Homo sapiens	Human GDF-3 (hGDF-3) polypeptide encoding cDNA.	1950	100
913	Y19757	Homo sapiens	SEQ ID NO 475 from WO9922243.	1361	100
914	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	112	48
915	U14971	Homo sapiens	ribosomal protein S9	886	90
916	AF172854	Homo sapiens	cardiotrophin-like cytokine CLC	1204	99
917	AC005525	Homo sapiens	F22162_1	1963	100
918	AF166350 Y87285	Homo sapiens	ST7 protein	4711 430	99
919		Homo sapiens	Human signal peptide containing protein HSPP- 62 SEQ ID NO:62.		100
920	Y36131	Homo sapiens	Human secreted protein #3.	465	88
921	AF193766	Homo sapiens	cytokine-like protein C17	724 357	100
922 923	Y95013 X75208	Homo sapiens Homo sapiens	Human secreted protein vc48_1, SEQ ID NO:66.  protein tyrosine kinaso-receptor	5256	100
923	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	813	98
925	AB039886	Homo sapiens	down-regulated in gastric cancer	785	78
926	G03368	Homo sapiens	Human secreted protein, SEQ ID NO: 7449.	55	50
927	Y48606	Homo sapiens	Human breast tumour-associated protein 67.	539	100
928	Y36151	Homo sapiens	Human secreted protein #23.	668	100
929	AF110399	Homo sapiens	elongation factor Ts	1666	100
930	AF210317	Homo sapiens	facilitative glucose transporter family member GLUT9	2763	99
931	Y73328	Homo sapiens	HTRM clone 082843 protein sequence.	931	100
932	G01959	Homo sapiens	Human secreted protein, SEQ ID NO: 6040.	274	100
933	U47924	Homo sapiens	B-cell receptor associated protein	1469	100
934	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	529	93
935	AB039371	Homo sapiens	mitochondrial ABC transporter 3	196	63
936	X56385	Canis familiaris	rab8	1064	100
937	B08906	Homo sapiens	Human secreted protein sequence encoded by gene 16 SEQ ID NO:63.	117	44
938	M13692	Homo sapiens	alpha-1 acid glycoprotein precursor	1064	99
939	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	515	42
940	Y16630	Homo sapiens	Human Putative Adrenomedullin Receptor (PAR).	1904	99
941	AC005102	Homo sapiens	small inducible cytokine subfamily A member 24	627	99
942	M12886	Homo sapiens	T-cell receptor beta chain	1289	81
943	AF226046	Homo sapiens	GK003	1049	98
944	Y36078	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 463.	667	100
945	M22877	Homo sapiens	cytochrome c	565	100
946	W67869	Homo sapiens	Human secreted protein encoded by gene 63 clone HHGDB72.	551	93
947	W67859	Homo sapiens	Human secreted protein encoded by gene 53 clone HBMCL41.	283	100
948	W85726	Homo sapiens	Novel protein (Clone BG33 7).	789	100
949	AJ242015	Homo sapiens	eMDC II protein	4236	100
950	G04075	Homo sapiens	Human secreted protein, SEQ ID NO: 8156.	567	99

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	140.			Score	Identity
951	AF110645	Homo sapiens	candidate tumor suppressor p33 ING1 homolog	1314	100
952	Y36111	Homo sapiens	Extended human secreted protein sequence, SEO	402	70
752	130111	monio sapiens	ID NO. 496.	402	/0
953	AB012109	Homo sapiens	APC10	990	100
954	AF246221	Homo sapiens	transmembrane protein BRI	1405	100
955	AF054986	Homo sapiens	putative transmembrane GTPase	1883	100
956	W74726	Homo sapiens			
957	Y27096		Human secreted protein fg949_3.	1879	100
957 958	AJ222967	Homo sapiens	Human viral receptor protein (ACVRP).  cystinosin	1581	100
		Homo sapiens		1920	100
959	Y53052	Homo sapiens	Human secreted protein clone df202_3 protein sequence SEQ ID NO:110.	587	100
960	G02694	Homo sapiens	Human secreted protein, SEQ ID NO: 6775.	283	100
961	AF151855	Homo sapiens	CG1-97 protein	1214	96
962	U26592	Homo sapiens	diabetes mellitus type I autoantigen	250	65
963	AL050306	Homo sapiens	dJ475B7.2 (novel protein)	3796	100
964	AF078859	Homo sapiens	PTD004	2089	100
965	AB020315	Homo sapiens	homologue of mouse dkk-1 gene:Acc# AF030433	1466	100
966	X04571	Homo sapiens	precursor polypeptide (AA -22 to 1185)	6580	99
967	AF146019	Homo sapiens	hepatocellular carcinoma antigen gene 520	993	99
968	AF071002	Homo sapiens	minK-related peptide 1; MiRP1	632	100
969	AB021227	Homo sapiens	membrane-type-5 matrix metalloproteinase	3545	100
970	AF180920	Homo sapiens	cyclin L ania-6a	1579	100
971	AF105365	Homo sapiens	K-Cl cotransporter KCC4	5621	99
972	AF083248	Homo sapiens			
			ribosomal protein L26 homolog	739	100
973	AJ132429	Homo sapiens	hyperpolarization-activated cyclic nucleotide gated cation channel hHCN4	6295	100
974	W61619	Homo sapiens	Clone HTPEF86 of TM4SF superfamily.	454	100
975	AF155100	Homo sapiens	zinc finger protein NY-REN-21 antigen	2261	100
976	AF275948	Homo sapiens	ABCA1	11763	99
977	AB026891	Homo sapiens	cystine/glutamate transporter	2552	100
978	AF117657	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP80	3348	99
979	AF044201	Rattus norvegicus	neural membrane protein 35; NMP35	1570	92
980	AF119297	Homo sapiens	neuroendocrine-specific protein-like protein 1	1170	99
981	AF155652	Homo sapiens	potassium channel modulatory factor	1983	99
982	W88499	Homo sapiens	Human stomach carcinoma clone HP10412- encoded protein.	1553	99
983	Z56281	Homo saplens	interferon regulatory factor 3	2012	98
984	AB026125	Homo sapiens	ART-4	2160	100
985	Y14482	Homo sapiens	Fragment of human secreted protein encoded by	172	70
			gene 17.		
986	AB023888	Homo sapiens	b-chemokine receptor CCR4	1895	100
987	W27291	Homo sapiens	Human H1075-1 secreted protein 5' end.	712	100
988	AF153450	Manduca sexta	juvenile hormone esterase binding protein	226	32
989	G03697	Homo sapiens	Human secreted protein, SEQ ID NO: 7778.	194	88
990	AF204159	Homo sapiens	potassium large conductance calcium-activated channel beta 3a subunit	1486	100
991	G02061	Homo sapiens	Human secreted protein, SEQ ID NO: 6142.	558	99
992	AL031266	Caenorhabditi s elegans	VM106R.1	327	40
993	Y66749	Homo sapiens	Membrane-bound protein PRO1124.	4730	99
994	G01246				
		Homo sapiens	Human secreted protein, SEQ ID NO: 5327.	141	77
995	AF133845	Homo sapiens	corin	5811	99
996	AF117756	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP150	4999	100
97	W62066	Homo sapiens	Human stem cell antigen 2.	284	93
998	Y87173	Homo sapiens	Human secreted protein sequence SEQ ID NO:212.	725	100
999	Y13379	Homo sapiens	Amino acid sequence of protein PRO263.	1654	99
1000	Y95008	Homo sapiens	Human secreted protein vf3_1, SEQ ID NO:56.	676	47
1001	AF190167	Homo sapiens	membrane associated protein SLP-2	1747	100

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1002	G01234	Homo sapiens	Human secreted protein, SEQ ID NO: 5315.	398	96
1003	W73420	Homo sapiens	Human secreted protein encoded by Gene No. 24.	2150	100
1004	X12791	Homo sapiens	19kD SRP-protein (AA 1 - 144)	742	100
1005	M23323	Homo sapiens	membrane protein	642	100
1006	X63745	Homo sapiens	KDEL receptor	326	98
1007	Y35997	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 382.	824	99
1008	AB032918	Hylobates moloch	dopamine receptor D4	92	35
1009	Y91680	Homo sapiens	Human secreted protein sequence encoded by gene 81 SEQ ID NO:353.	1372	99
1010 ·	AL136125	Homo sapiens	dJ304B14.1 (novel protein)	825	98
1011	G03733	Homo sapiens	Human secreted protein, SEQ ID NO: 7814.	379	98
1012	Y17531	Homo sapiens	Human secreted protein clone BL205 14 protein.	818	97
1013	G00724	Homo sapiens	Human secreted protein, SEQ ID NO: 4805.	462	100
1014	AF288092	Naegleria gruberi	haem lyase	114	37
1015	AB045292	Homo sapiens	M83 protein	3867	99
1016	X15940	Homo sapiens	ribosomal protein L31 (AA 1-125)	644	100
1017	Y94873	Homo sapiens	Human protein clone HP02632.	1876	100
1018	AL024498	Homo sapiens	dJ417M14.1 (novel protein)	589	100
1019	X83425	Homo sapiens	Lutheran blood group glycoprotein	3054	99
1020	W03516	Homo sapiens	Prostaglandin DP receptor.	1864	100
1021	G03960	Homo sapiens	Human secreted protein, SEQ ID NO: 8041.	398	100
1022	Y91689	Homo sapiens	Human secreted protein sequence encoded by genc 93 SEQ ID NO:362.	768	100
1023	AE000660	Homo sapiens	hADV36S1	573	100
1024	AF132965	Homo sapiens	CG1-31 protein	1550	100
1025	W92380	Homo sapiens	Human TR-interacting protein S103a.	1466	97
1026	R66278	Homo sapiens	Therapeutic polypeptide from glioblastoma cell line.	830	100
1027	X65614	Homo sapiens	S100P calcium-binding protein	476	100
1028	Y41741	Homo sapiens	Human PRO704 protein sequence.	1323	100
1029	AJ001014	Homo sapiens	RAMP1	806	100
1030	W63682	Homo sapiens	Human secreted protein 2.	1354	99
1031	AK023007	Homo sapiens	unnamed protein product	766	100
1032 1033	W97900 Y82453	Homo sapiens	Human SR-BI class B scavenger.	2672	99
		Homo sapiens	Human TGC-440 secretory protein SEQ ID NO:1.	639	99
1034	Y73473	Homo sapiens	Human secreted protein clone yd178_1 protein sequence SEQ ID NO:168.	752	93
1035	Y86468	Homo sapiens	Human gene 48-encoded protein fragment, SEQ ID NO:383.	96	90
1036 1037	U09813 AJ242832	Homo saplens	mitochondrial ATP synthase subunit 9 precursor	698	100
1037	X66403	Homo sapiens	calpain acetylcholine receptor epsilon subunit CHRNE	3699	99
1038	AJ242730	Homo sapiens Homo sapiens	polyhomeotic 2	2574	100
1040	AF169968	Mus musculus	DNA binding protein DESRT	1310 1453	80
1041	X52563	Bos taurus	permability increasing protein	383	29
1041	G00368	Homo sapiens	Human secreted protein. SEO ID NO: 4449.	75	50
1043	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	60	53
1043	M94582	Homo sapiens	interleukin 8 receptor B	1850	100
1045	AL080239	Homo sapiens	bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunlt))	1704	50
1046	AF125101	Homo sapiens	HSPC040 protein	580	100
1047	W74809	Homo sapiens	Human secreted protein encoded by gene 81 clone HMWDN32.	176	100
1048	AL022238	Homo sapiens	dJ1042K10.4 (novel protein)	2201	100
1049 ·	W88667	Homo sapiens	Secreted protein encoded by gene 134 clone HAIBP89.	1559	99
1050	AF097518	Homo sapiens	liver-specific transporter	2820	100

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1051	W78324	Homo sapiens	Fragment of human secreted protein encoded by gene 81.	1318	98
1052	Y21851	Homo sapiens	Human signal peptide-contianing protein (SIGP) (clone ID 2328134).	1643	95
1053	AL163815	Arabidopsis thaliana	putative protein	661	62
1054	Y76200	Homo sapiens	Human secreted protein encoded by gene 77.	262	100
1055	AJ276567	Homo sapiens	TC10-like Rho GTPase	1160	100
1056	Y27620	Homo sapiens	Human secreted protein encoded by gene No. 54.	154	96
1057	D14530	Homo sapiens	ribosomal protein	745	100
1058	AF132000	Homo sapiens	TADA1 protein	1132	100
1059	AL031778	Homo sapiens	dJ34B21.1 (novel BZRP (benzodiazzpine receptor (peripheral) (MBR, PBR, PBKS, IBP, lsoquinoline-binding protein)) LIKE protein)	920	100
1060	AF227135	Homo sapiens	candidate taste receptor T2R9	134	33
1061	Y27575	Homo sapiens	Human secreted protein encoded by gene No. 9.	1392	100
1062 1063	Z11697 AF123757	Homo sapiens	HB15	1088	100
1064	AF123757 AF155135	Homo sapiens Homo sapiens	putative transmembrane protein	819	100
1065	Y41674	Homo sapiens	novel retinal pigment epithelial cell protein Human channel-related molecule HCRM-2.	2932	99
1066	AJ250042	Homo sapiens	Rab5 GDP/GTP exchange factor homologue	936 2575	99 100
1067	Y36087	Homo sapiens	Extended human secreted protein sequence, SEQ	770	85
1068	Y94959	Homo sapiens	ID NO. 472.  Human secreted protein clone mc300 1 protein	301	100
1069	Y94959	Homo sapiens	sequence SEQ ID NO:124.  Human secreted protein clone mc300 1 protein	301	100
1070	W64535	Homo sapiens	sequence SEQ ID NO:124.  Human leukocyte cell clone HP00804 protein.	2014	99
1071	X03145	Homo sapiens	pot. ORF III	148	50
1072	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	821	91
1073	X82200	Homo sapiens	gpStaf50	249	62
1074	G03213	Homo sapiens	Human secreted protein, SEQ ID NO: 7294.	99	47
1075	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	506	55
1076	G03187	Homo sapiens	Human secreted protein, SEQ ID NO: 7268.	424	98
1077	L25899	Homo sapiens	ribosomal protein L10	332	76
1078	Y91447	Homo sapiens	Human secreted protein sequence encoded by gene 48 SEQ ID NO:168.	898	97
1079	G01862	Homo sapiens	Human secreted protein, SEQ ID NO: 5943.	290	89
1080 · 1081	AB039723	Homo sapiens	WNT receptor frizzled-3	1376	92
1082	AB020527 L13802	Homo sapiens Homo sapiens	Na/PO4 cotransporter homolog ribosmal protein small subunit	269	100
1083	W75098	Homo sapiens	Human secreted protein encoded by gene 42	499 143	80
1084	G03564		clone HSXBI25.		81
1085	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	83	51
1086	AF090942	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 8144. PRO0657	88 124	43 64
1087	G00517	Homo sapiens	Human secreted protein, SEO ID NO: 4598.	129	41
1088	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	126	36
1089	AF140631	Homo sapiens	G-protein coupled receptor 14	364	82
1090	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	114	32
1091	S72304	Mus sp.	LMW G-protein	146 .	83
1092	W88708	Homo sapiens	Secreted protein encoded by gene 175 clone HEMAM41.	405	100
1093	W85612	Homo sapiens	Secreted protein clone fh123_5.	4358	97
1094	Y53012	Homo sapiens	Human secreted protein clone pm514_4 protein sequence SEQ ID NO:30.	1013	99
1095	Y92345	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-62.	409	100
1096	AF090942	Homo sapiens	PRO0657	147	60
1097	L24521	Homo sapiens	transformation-related protein	166	58
1098	X56932	Homo sapiens	23 kD highly basic protein	490	70
1099	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	83	35
1100	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	149	59

SEQ	Accession	Species	Description	Smith-	1%
ID T	No.			Waterman	Identity
NO:	j			Score	
1101	AF119851	Homo sapiens	PRO1722	183	72
1102	G04086	Homo sapiens	Human secreted protein, SEQ ID NO: 8167.	207	62
1103	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	91	52
1104	X74856	Mus	ribosomal protein L28	128	69
		musculus			1
1105	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	130	62
1106	G03133	Homo sapiens	Human secreted protein, SEQ ID NO: 7214.	122	48
1107	G03040	Homo sapiens	Human secreted protein, SEQ ID NO: 7121.	69	43
1108	AF039942	Homo sapiens	HCF-binding transcription factor Zhangfei	744	99
1109	AF201951	Homo sapiens	high affinity immunoglobulin epsilon receptor	738	94
_	ł	1	beta subunit	Į -	1
1110	AF111108	Mus	transient receptor potential 2	223	79
	<u> </u>	musculus	· ·		
1111	AF119900	Homo sapiens	PRO2822	144	59
1112	Y16589	Homo sapiens	A protein that interacts with presentlins.	265	39
1113	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	178	67
1114	Y02999	Homo sapiens	Fragment of human secreted protein encoded by	164	63
			gene 121.	1	
1115	Y30811	Homo sapiens	Human secreted protein encoded from gene 1.	1217	99
1116	X51394	Xenopus	APEG precursor protein	130	40
		laevis	·		
1117	M27826	Homo sapiens	neutral protease large subunit	442	65
1118	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	72	60
1119	G03602	Homo sapiens	Human secreted protein, SEQ ID NO: 7683.	491	97
1120	Y35906	Homo sapiens	Extended human secreted protein sequence, SEQ	244	97
	ļ	<u> </u>	ID NO. 155.		
1121	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	122	65
1122	Y00337	Homo sapiens	Human secreted protein encoded by gene 81.	110	90
1123	AF084830	Homo sapiens	two pore domain K+ channel; TASK-2	703	94
1124	AF212862	Homo sapiens	membrane interacting protein of RGS16	442	88
1125	W64469	Homo sapiens	Human secreted protein from clone CW795 2.	191	53
1126	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	154	100
1127	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	165	100
1128	Y84320	Homo sapiens	Human cardiovascular system associated protein	815	99
1100	000105	<del> </del>	kinase-1.		
1129	G02105	Homo sapiens	Human secreted protein, SEQ ID NO: 6186.	88	73
1130	Y32923	Homo sapiens	Transmembrane domain containing protein clone HP01512.	700	100
1131	Y29817			-	<u> </u>
1132	Y91644	Homo sapiens	Human synapse related glycoprotein 2.	260	91
1132	191044	Homo sapiens	Human secreted protein sequence encoded by	525	96
1133	Y91449	Illama appiana	gene 43 SEQ ID NO:317.		100
1133	171449	Homo sapiens	Human secreted protein sequence encoded by gene 49.SEO ID NO:170.	542	100
1134	AB017908 ·	Homo sapiens	4F2 light chain	2200	02
1135	X51760	Homo sapiens	zinc finger protein (583 AA)	2399	93
1136	Y99426	Homo sapiens	Human PRO1604 (UNQ785) amino acid	312	55
1130	177420	1 10mo sapiciis	sequence SEO ID NO:308.	917	72
1137	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	102	50
1138	AF155106	Homo sapiens	NY-REN-36 antigen	768	91 ·
1139	AL031055	Homo sapiens	dJ28H20.1 (novel protein similar to membrane	117	50
7	[ 2 3202 1022	I TYANIA SUPTERS	mentale in mentales	117	<i>3</i> 0
	j	) 1	transport proteins)		
1140	AF011359	Bos taums	transport proteins)	138	96
1140 1141	AF011359 Y70018	Bos taurus Homo saniens	regulator of G-protein signaling 7	138	96
1140 1141	AF011359 Y70018	Bos taurus Homo sapiens	regulator of G-protein signaling 7 Human Protease and associated protein-12	138 623	96 100
1141	Y70018	Homo sapiens	regulator of G-protein signaling 7 Human Protease and associated protein-12 (PPRG-12).	623	100
1141 1142	Y70018 G04091	Homo sapiens Homo sapiens	regulator of G-protein signaling 7 Human Protease and associated protein-12 (PPRG-12). Human secreted protein, SEQ ID NO: 8172.	623 113	100 38
1141	Y70018	Homo sapiens Homo sapiens Canis	regulator of G-protein signaling 7 Human Protease and associated protein-12 (PPRG-12).	623	100
1141 1142 1143	Y70018 G04091 AB030235	Homo sapiens Homo sapiens Canis familiaris	regulator of G-protein signaling 7  Human Protease and associated protein-12 (PPRG-12).  Human secreted protein, SEQ ID NO: 8172.  D4 dopamine receptor	623 113 89	100 38 48
1141 1142	Y70018 G04091	Homo sapiens Homo sapiens Canis	regulator of G-protein signaling 7  Human Protease and associated protein-12 (PPRG-12).  Human secreted protein, SEQ ID NO: 8172.  D4 dopamine receptor  Human secreted protein clone pv6_1 protein	623 113	100 38
1141 1142 1143 1144	Y70018 G04091 AB030235 Y94922	Homo sapiens Homo sapiens Canis familiaris Homo sapiens	regulator of G-protein signaling 7  Human Protease and associated protein-12 (PPRG-12).  Human secreted protein, SEQ ID NO: 8172.  D4 dopamine receptor  Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50.	623 113 89 539	38 48 88
1141 1142 1143 1144 1145	Y70018 G04091 AB030235 Y94922 X99962	Homo sapiens Homo sapiens Canis familiaris Homo sapiens Homo sapiens	regulator of G-protein signaling 7  Human Protease and associated protein-12 (PPRG-12).  Human secreted protein, SEQ ID NO: 8172.  D4 dopamine receptor  Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50.  rab-related GTP-binding protein	623 113 89 539	38 48 88
1141 1142 1143 1144 1145 1146	Y70018 G04091 AB030235 Y94922 X99962 G03807	Homo sapiens  Homo sapiens Canis familiaris Homo sapiens Homo sapiens Homo sapiens	regulator of G-protein signaling 7  Human Protease and associated protein-12 (PPRG-12).  Human secreted protein, SEQ ID NO: 8172.  D4 dopamine receptor  Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50.  rab-related GTP-binding protein  Human secreted protein, SEQ ID NO: 7888.	623 113 89 539 398 168	38 48 88 96 79
1141 1142 1143 1144 1145	Y70018 G04091 AB030235 Y94922 X99962	Homo sapiens Homo sapiens Canis familiaris Homo sapiens Homo sapiens	regulator of G-protein signaling 7  Human Protease and associated protein-12 (PPRG-12).  Human secreted protein, SEQ ID NO: 8172.  D4 dopamine receptor  Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50.  rab-related GTP-binding protein	623 113 89 539	38 48 88

SEQ	Accession	Species	Description	Smith-	1%
ID	No.			Waterman	Identity
NO:	ļ		-	Score	<u> </u>
1150	C02420	s elegans	cerevisiae zinc resistance protein		
1150	G03438	Homo sapiens	Human secreted protein, SEQ ID NO: 7519.	117	62
1151	G01003	Homo sapiens	Human secreted protein, SEQ ID NO: 5084.	181,	80
1152	G03798 X88799 *	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	198	63
1154	D85245	Oryza sativa	DNA binding protein	95	41
1155	R74272	Homo sapiens	TR3beta	155	96
1156	Y86265	Homo sapiens	Turnour suppressor protein, p53.	341	87
L		Homo sapiens	Human secreted protein HUSXE77, SEQ ID NO:180.	99	41
1157	G02577	Homo sapiens	Human secreted protein, SEQ ID NO: 6658.	263	98
1158	AF104334	Homo sapiens	putative organic anion transporter	185	42
1159	G01393	Homo sapiens	Human secreted protein, SEQ ID NO: 5474.	173	57
1160	W75771	Homo sapiens	Human GTP binding protein APD08.	224	81
1161	AF216833	Homo sapiens	M-ABC2 protein	410	83
1162	W67816	Homo sapiens	Human secreted protein encoded by gene 10 clone HCEMU42.	1156	100
1163	AF119851	Homo sapiens	PRO1722	230	70
1164	Y87252	Homo sapiens	Human signal peptide containing protein HSPP- 29 SEQ ID NO:29.	113	31
1165	W64537	Homo sapiens	Human liver cell clone HP01148 protein.	338	82
1166	AF269286	Homo sapiens	HC6	134	64
1167	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	149	51
1168	D90789	Escherichia coli	Dipeptide transport system permease protein DppC.	411	90
1169	R63783	Homo sapiens	TG0847 protein.	344	90
1170	Y45274	Homo sapiens	Human secreted protein encoded from gene 18.	478	98
1171	D64154	Homo sapiens	Mr 110,000 antigen	347	96
1172	AB026256	Homo sapiens	organic anion transporter OATP-B	311	67
1173	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	60	52
1174	D87717	Homo sapiens	similar to human GTPase-activating protein(A49869)	178	59
1175	M64716	Homo sapiens	ribosomal protein	391	78
1176	R08330	Homo sapiens	Human IL-7 receptor clone H6.	285	67
1177	L06505	Homo sapiens	ribosomal protein L12	242	72
1178	AJ251885	Homo sapiens	organic cation transporter (OCT2)	276	88
1179	G03258	Homo saplens	Human secreted protein, SEQ ID NO: 7339.	155	71
1180	G01207	Homo sapiens	Human secreted protein, SEQ ID NO: 5288.	282	90
1181	AF181856	Rattus	tRNA selenocysteine associated protein	249	62
		norvegicus	•	,	
1182	AF161524	Homo sapiens	HSPC176	138	90
1183	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	282	66
1184	Y02671	Homo sapiens	Human secreted protein encoded by gene 22 clone HMSJW18.	107	71 .
	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	88	69
1186	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	118	46
1187	AB032905	Hylobates concolor	dopamine receptor D4	96	37
1188	G00956	Homo sapiens	Human secreted protein, SEQ ID NO: 5037.	292	78
1189	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	178	79
1190	G03361	Homo sapiens	Human secreted protein, SEQ ID NO: 7442.	324	76
1191	AF117755	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP230	187	70
1192	Y70455	Homo sapiens	Human membrane channel protein-5 (MECHP-5).	202	67
1193	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	99	42
1194	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	192	76
1195	W29661	Homo sapiens	Homo sapiens C1542 2 clone secreted protein.	2001	98
1196	Y14104	Homo sapiens	Human GABAB receptor 1d protein sequence.	239	69
1197	X61972	Homo sapiens	macropain subunit iota	149	90
1198	G00534	Homo sapiens	Human secreted protein, SEQ ID NO: 4615.	145	51
1199	Y86260	Homo sapiens	Human secreted protein HELHN47, SEQ ID	1089	89
1200	G02607	Homo sapiens	NO:175. Human secreted protein, SEQ ID NO: 6688.	154	57
			2. The sources protein, SEQ ID NO. 0000.	1,54	<u> </u>

SEQ	Accession	Species	Description	Smith-	%
ID NO:	No.		·	Waterman Score	Identity
1201	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	404	50
1202	M27826	Homo sapiens	neutral protease large subunit	202	49
1203	Y73424	Homo sapiens	Human secreted protein clone yi4_1 protein	265	61
	<u> </u>		sequence SEQ ID NO:70.		
1204	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	625	98
1205	Y36203	Homo sapiens	Human secreted protein #75.	219	59
1206	U78111	Gallus gallus	AQ	205	57
1207	AF095448	Homo sapiens	putative G protein-coupled receptor	416	76
1208	AF116715	Homo sapiens	PRO2829	127	75
1209	AF099137	Homo sapiens	MaxiK channel beta 2 subunit	475	95
1210	AF205718	Homo sapiens	hepatocellular carcinoma-related putative tumor suppressor	423	79
1211	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	224	70
1212	G00719	Homo sapiens	Human secreted protein, SEQ ID NO: 4800.	117	44
1213	G01009	Homo sapiens	Human secreted protein, SEQ ID NO: 5090.	351	73
1214	AF090942	Homo sapiens	PRO0657	124	70
1215	Y14427	Homo sapiens	Human secreted protein encoded by gene 17	99	77
1216	G03905	Homo sapiens	clone HSIEA14.  Human secreted protein, SEQ ID NO: 7986.	173	57
1217	Y57897			1173	100
1217	J00194	Homo sapiens	Human transmembrane protein HTMPN-21.	454	78
1219	Y59709	Homo sapiens	hla-dr antigen alpha chain		92
1220	W81576	Homo sapiens Homo sapiens	Secreted protein 76-28-3-A12-FL1.	725	
	ļ		EBV-induced G-protein coupled receptor (EBI-2) polypeptide.		100
1221	W96745	Homo sapiens	High affinity immunoglobulin E receptor-like protein (IGERB).	650	98
1222	Y35911	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 160.	135	31
1223	Y00278	Homo sapiens	Human secreted protein encoded by gene 21.	260	95
1224	AF161422	Homo sapiens	HSPC304	568	90
1225	U14970	Homo sapiens	ribosomal protein S5	202	95
1226	G01733	Homo sapiens	Human secreted protein, SEQ ID NO: 5814.	610	100
1227	AF099973	Mus musculus	schlafen2	333	56
1228	G01218	Homo sapiens	Human secreted protein, SEQ ID NO: 5299.	155	81
1229	AF217188	Mus musculus	YIPIB	801	63
1230	AF176813	Homo sapiens	soluble adenylyl cyclase	275	100
1231	X98333	Homo sapiens	organic cation transporter	1704	100
1232	W74955	Homo sapiens	Human secreted protein encoded by gene 77 clone HOEAS24.	212	53
1233	Y94940	Homo sapiens	Human secreted protein clone yi62_1 protein sequence SEQ ID NO:86.	526	100
1234	U76618	Mus musculus	N-RAP	482	82
1235	AF044924	Homo sapiens	hook2 protein	380	97
1236	G01459	Homo sapiens	Human secreted protein, SEQ ID NO: 5540.	417	100
1237	AF000018	Homo sapiens	adapter protein	164	84
1238	W88633	Homo sapiens	Secreted protein encoded by gene 100 clone HE8EU04.	250	90 .
1239	W29660	Homo sapiens	Homo sapiens CH27 1 clone secreted protein.	697	98
1240	AF004161	Oryctolagus	peroxisomal Ca-dependent solute carrier	154	52
1041	¥/00/110	cuniculus		1	100
1241	Y92710	Homo sapiens	Human membrane associated protein Zsig24.	709	97
1242	Y95002	Homo sapiens	Human secreted protein vc34_1, SEQ ID NO:44.	908	88
1243	Y44905	Homo sapiens	Human potassium channel molecule ERG-LP2 partial protein.	325	100
1244	AF284422	Homo sapiens	cation-chloride cotransporter-interacting protein	511	97
1245	Y53629	Homo sapiens	A bone marrow secreted protein designated BMS115.	1888	93
1246	AB039371	Homo sapiens	mitochondrial ABC transporter 3	389	97
		Homo sapiens	Extended human secreted protein sequence, SEQ	168	39

SEQ	Accession	Species	Description	Smith-	1%
ID	No.			Waterman	Identity
NO:		<b></b> _	<u> </u>	Score	<u> </u>
		<u> </u>	ID NO. 160.		ļ
1248	AF072509	Rattus	glutamate receptor interacting protein 2	559	90
1249	AF247042	Homo sapiens		<del> </del>	100
1249	B08974	Homo sapiens	tandem pore domain potassium channel TRAAK	661	98
			Human secreted protein sequence encoded by gene 27 SEQ ID NO:131.	1087	97
1251	L15313	Caenorhabditi s elegans	putative	858	59
1252	Y29338	Homo sapiens	Human secreted protein clone it217_2 alternate reading frame protein.	278	75
1253	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	211	92
1254	G03074	Homo sapiens	Human secreted protein, SEQ ID NO: 7155.	294	83
1255	G01818	Homo sapiens	Human secreted protein, SEQ ID NO: 5899.	253	91
1256	AF286368	Homo sapiens	eppin-l	222	54
1257	AF220264	Homo sapiens	MOST-1	87	93
1258	G02227	Homo sapiens	Human secreted protein, SEQ ID NO: 6308.	281	78
1259	Y07970	Homo sapiens	Human secreted protein fragment #2 encoded from gene 26.	81	94
1260	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain ligand (clone 3TW).	986	100
1261	AF140674	Homo sapiens	zinc metalloprotease ADAMTS6	172	36
1262	U28369	Homo sapiens	semaphorin V	237	67
1263	Y07049	Homo sapiens	Renal cancer associated antigen precursor sequence.	288	71
1264	Y36153	Homo sapiens	Hurnan secreted protein #25.	187	80
1265	Y78114	Homo sapiens	Human cytokine signal regulator CKSR-2 SEQ ID NO:2.	723	93
1266	Y13397	Homo sapiens	Amino acid sequence of protein PRO334.	191	100
1267	AF030558	Rattus norvegicus	phosphatidylinositol 5-phosphate 4-kinase gamma	859	95
1268	U73167	Homo sapiens	candidate tumor suppressor gene LUCA-1	159	96
1269	AF190664	Mus musculus	LMBR2	552	76
1270	AL050332	Homo sapiens	dJ570F3.1 (homolog of the rat synaptic ras GTPase-activating protein p135 SynGAP)	820	98
1271	G02126	Homo sapiens	Human secreted protein, SEQ ID NO: 6207.	131	95
1272	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	253	92
1273	AL035661	Homo sapiens	dJ568C11.3 (novel AMP-binding enzyme similar to acetyl-coenzyme A synthethase (acetate-coA ligase))	1280	100
1274	AF064748	Mus musculus	S3-12	3523	61
1275	D17554	Homo sapiens	TAXREB107	377	78
1276	Y30715	Homo sapiens	Amino acid sequence of a human secreted protein.	643	90
1277	AF146760	Homo sapiens	septin 2-like cell division control protein	707	100
1278	Y05069	Homo sapiens	Human PIGR-2 protein sequence.	281	46
1279	X59668	Oryctolagus cuniculus	aorta CNG channel (rACNG)	267	85
1280	G01051	Homo sapiens	Human secreted protein, SEQ ID NO: 5132.	489	98
1281	G03411	Homo sapiens	Human secreted protein, SEQ ID NO: 7492.	120	43
1282	AF055084	Homo sapiens	very large G-protein coupled receptor-1	1635	100
1283	AF117814	Mus musculus	odd-skipped related 1 protein	357	98
1284	U87318	Xenopus laevis	NaDC-2	535	60
1285	AF061346	Mus musculus	Edp1 protein	452	68
1286	AB030182	Mus musculus	contains transmembrane (TM) region	582	68
1287	A13595	synthetic	immunosuppresive protein PP15	185	97
1288	AF254411	Homo sapiens	ser/arg-rich pre-mRNA splicing factor SR-A1	837	100
1289	AF084205	Rattus	serine/threonine protein kinase TAO1	319	98
	1 30.203	norvegicus	owner an action of desired the terms of the	ر ا	ا آ

SEQ ID No:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1290	AF038563	Homo sapiens	membrane associated guanylate kinase 2	523	100
1291	AF034837	Homo sapiens	double-stranded RNA specific adenosine deaminase	468	100
1292	M15888	Bos taurus	endozepine-related protein precursor	937	87
1293	AB010692	Arabidopsis thaliana	ATP-dependent RNA helicase-like protein	636	45
1294	AF209923	Homo sapiens	orphan G-protein coupled receptor	1570	100
1295	W67828	Homo sapiens	Human secreted protein encoded by gene 22 clone HFEAF41.	504	98
1296	AC004832	Homo sapiens	similar to 45 kDa secretory protein; similar to CAA10644.1 (PID:g4164418)	648	65
1297	X80035	Oryctolagus cuniculus	cysteine rich hair keratin associated protein	575	70
1298	G02645	Homo sapiens	Human secreted protein, SEQ ID NO: 6726.	223	97
1299	Y59440	Homo sapiens	Human delta3 fragment #4.	122	32
1300	W70504	Homo sapiens	Leukocyte seven times membrane-penetrating type receptor protein JEG18.	459	81
1301	Y67315	Homo sapiens	Human secreted protein BL89_13 amino acid sequence.	3916	99
1302	M77693	Homo sapiens	spermidine/spermine N1-acety/transferase	174	96
1303	G01331	Homo sapiens	Human secreted protein, SEQ ID NO: 5412.	254	69
1304	G01491	Homo sapiens	Human secreted protein, SEQ ID NO: 5572.	747	99
1305	AF148509	Homo sapiens	alpha 1,2-mannosidase	602	98
1306	G01658	Homo sapiens	Human secreted protein, SEQ ID NO: 5739.	333	98
1307	Y90899	Homo sapiens	D1-like dopamine receptor activity modifying protein SEQ ID NO:1.	332	98
1308	AF033120	Homo sapiens	p53 regulated PA26-T2 nuclear protein	348	52
1309	Y73388	Homo sapiens	HTRM clone 3376404 protein sequence.	147	66
1310 1311	AF063243 AF224494	Bos taurus	ribosomal protein L30	296	90
	<u> </u>	Mus musculus	arsenite inducible RNA associated protein	688	70
1312	Y73342	Homo sapiens	HTRM clone 2709055 protein sequence.	1154	100
1313	Y99419	Homo sapiens	Human PRO1780 (UNQ842) amino acid sequence SEQ ID NO:282.	1145	78
1314 1315	AF116667	Homo sapiens	PRO1777	433	97
	W75100	Homo sapiens	Human secreted protein encoded by gene 44 clone HE8CJ26.	807	97
1316	AJ272078	Homo sapiens	APOBEC-1 stimulating protein	789	100
1318	AB041533 U19617	Homo sapiens	sperm antigen Elf-1	2607	98
1319		Mus musculus		806	92
	U82598	Escherichia coli	ferric enterobactin transport protein	768	100
1320	D90892 W67847	Escherichia coli  Homo sapiens	SORBITOL-6-PHOSPHATE 2- DEHYDROGENASE (EC 1.1.1.140) (GLUCITOL-6-PHOSPHATE DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE).	709	100
	1		Human secreted protein encoded by gene 41 clone HPBCJ74.	601	92
1322 1323	AJ276101 AJ276101	Homo sapiens	GPRCSB protein	466	93
1324	Y58628	Homo sapiens Homo sapiens	GPRC5B protein	504	97
1325	U91561	Rattus norvegicus	Protein regulating gene expression PRGE-21.  pyridoxine 5'-phosphate oxidase	1584 1277	100 89
1326	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	1606	100
1327	Y32206	Homo sapiens	Human receptor molecule (REC) encoded by Incyte clone 2825826.	1531	90
1328	AF151048	Homo sapiens	HSPC214	657	85
1329	Y10530	Homo sapiens	olfactory receptor	1645	100
1330	AF180681	Homo sapiens	guanine nucleotide exchange factor	4314	99
1331	AF111856	Homo sapiens	sodium dependent phosphate transporter isoform NaPi-3b	3591	99
1332	Y13583	Homo sapiens	G-protein coupled receptor	2171	100
1334					

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	ſ		Waterman	Identity
NO:			•	Score	
1334	Y25755	Homo sapiens	Human secreted protein encoded from gene 45.	1380	96
1335	AF152325	Homo sapiens	protocadherin gamma A5	4742	99
1336	X74070	Homo sapiens	transcription factor BTF3	639	81
1337	AF095927	Rattus norvegicus	protein phosphatase 2C	1931	95
1338	G03877	Homo sapiens	Human secreted protein, SEQ ID NO: 7958.	621	100
1339	AL008582	Homo sapiens	bK223H9.2 (ortholog of A. thaliana F23F1.8)	626	100
1340	X61615	Homo sapiens	leukemia inhibitory factor receptor	5820	99
1341	Y01519	Homo sapiens	A carcinogenesis-inhibiting protein.	7528	97
1342	AF207600	Homo sapiens	ethanolamine kinase	2372	100
1343	U54807	Rattus norvegicus	GTP-binding protein	1167	97
1344	AC020579	Arabidopsis thaliana	putative phosphoribosylformylglycinamidine synthase; 25509-29950	3283	51
1345	Y28576	Homo sapiens	Secreted peptide clone pe503_1.	944	100
1346	W74787	Homo sapiens	Human secreted protein encoded by gene 58 clone HHFHN61.	1171 .	100
1347	M55542	Homo sapiens	guanylate binding protein isoform 1	2636	87
1348	AF183428	Homo sapiens	28.4 kDa protein	1329	100
1349	U70669	Homo sapiens	Fas-ligand associated factor 3	167	24
1350	AF295530	Homo sapiens	cardiac voltage gated potassium channel modulatory subunit	562	99

## TABLE 3

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
1	1351	A	2	sequence 337	1	nucleotide insertion  TPSLIHQAPTPCPAGLWG/PPNGHYHGS*PGC HWPQAPHRA****GLLPPRWLGHGLPGGPAAP WAASQWVDGVAGRLPGPAWSWHASGAAPA OPGPL*LLVPGSSGLPDPRDP
2	1352	A	27	100	366	IRNSSIRPMKERETKLSAKHMITCSASYDIRGL QIETTYHHTPIRMAKIQKT/GHHQC**ECGAT GTLIHGWWGCKVVEPLGKTVWOIPK
3	1353	A	40	3	314	·HASAHASVVLKDNSELEQQLGATGAYRARA LELEAEVAEMROMLQLEHPFVNGADKLRPD SMYVHLNEL*QSLVENMLLTVVDTH\RTPI*R SCNYTLALILFL
4	1354	A	74		292	TASALFSCPDGGSLAGFAGRRASFHLECLKR QKDRGGDISQKTVLPLHLVHHQVAHTFGQAT VTCQQARQSPG*RTNPE/ALQWVLPVSDGWH VLPLP
5	1355	A	78	114	850	ENCRVASNLPGVFFSEDTAQSGSYMRISAHPP NAGGEVSNGPKRKLTLMLNFSLPSSGLNAGA FYALSTLLNRMVIWHYPGEEVNAGRIGLTIVI AGMLGAVISGIWLDRSKTYKETTLVVYIMDT GGAWWCYTFYLGTGDTCG*CFTTAG\TMGFF MTGYLPLGFEFAVEL\SYPESEGISSGLLNISA QVFGIIFTISQGQIIDNYGTKPGNIFLCVFLTLG AALTAFIKADLRRQKANKETLEN
6	1356	A	81	97	376	EWFSYMLGSNMSVYHSP*SLEPLCKVLSES*A YLRVPFIRILLNAR*IRKAYKRMSLEIKLLI/RE *CLFQEMGLSLQWLYSARGDFFRATSRL
7	1357	A	93	2	872	TLSSACLIGDAWKELTTVAGAVSNQLLVWYP ATALADNKPVAPDRRISGHVGIIFSMSYLESK GLLATASEDRSVRIWKGGDLRVPGGRVQNIG HCFGHSARVWQVKLLENYLISAGEDCVCLV

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	Correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	<b>[</b>	1	ĺ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ì	l	ł	peptide		/=possible nucleotide deletion, \=possible
<u> </u>			<u> </u>	sequence		nucleotide insertion
	1	1	<u> </u>			WSHEGEILQAFRGHQGRGIRALAAHERQAWV
	ł	l	ł		ł	ITGGDDSGIRLWHLVGRGYRGLG/DLGSLLQ   VP**ARYTQGCDSGWLLATAGSD*YRGPVSL
		}				*RRGQVLGAAARG*TFPVLLPAGGSSWSRGL
ļ	}	ł	ŀ	ļ		RIVCYGQWGRSCQGCPHQHSNCCCGPDPVS
		l				WEGAQLELGPAWL
8	1358	A	106	3	350	FSSLLSGRISTLRDETGAILIDGDPAACAPIIKF
				ľ		LLTEELHLRGVSIYVLRHEAQIYGITPL\VCAL
	i				]	LI/CRRL*SDSCMRAALNDRGLYQVLILDGLV
	]		[	[	ļ	QCLGFVDSDSRKMVSTLT
9	1359	A	115	49	186	QAWAIFKGKYKEGDTGGPAVWKTRLRCALN
	L	L		_		KSSEFNEGPERERMDV
10	1360	A	123	2	1249	KGCRTQEKVDRTEVIRTCINPVYSKLFTVDFY
1	ļ	ļ	ļ			FEEVQRLRFEVHDISSNHNGLKEADFLGGME
	ĺ	1	ĺ			CTLGQIVSQRKLSKSLLKHGNTAGKSSITVIA
			1			EELSGNDDYVELAFNARKLDDKDFFSKSDPF
	1	ł	ł	ł	1	LEIFRMNDDATQQLVHRTEVVMNNLSPAWK
			1		ļ	SFKVSVNSLCSGDPDRRLKCIVWDWDSNGK   UDEICEETSTEEKEN OF CAMERICAL CONTROL OF CO
	]	ļ	ļ			HDFIGEFTSTFKEMRGAMEGKQVQWECINPK   YKAKKKNYKNSGTVILNLCKIHKMHSFLDYI
1	}	}	1		·	MGGCQIQFTVAIDFTASNGDPRNSCSLHYIHP
						YQPNEYLKALVAVGEICQDYDSDKMFPAFGF
	}		ļ		}	GARIPPEYTDSHDFAINFNEDNPECAGIQGVV
			1		i	EAYQSCF\PKAPTFTGPTNICPHSSRKVAKFRR
'	Ì	1	1			SEGN*HQGRAFAIIFILVDPGQVGVYSQDMGP
			Í			DNPGGHFV
11	1361	A	147	614	9	ACARKQLLGRTVFIWFVGQLLGGELKGYSKT
		1	ĺ		1	NTTSSRPASSRG\TLSSSSSSSSSLTKDALPSSL
		Ì	<b>!</b>		Į.	KSDSTTTTSGLVFPFRSLCVNPAKSSVSESVSSI
İ	ĺ	1 .	İ		İ	KILLSSSVKYLE*KRTSCCFPDSSESKLSQLSS
.			[			DERVSMGTSSRKPTNSSSSLGALKMSATS\*G
			Ì			SGSESPTPFFLTGLQSPPSTRPREPGLTTARNS   TTLTRDC
12	1362	A	177	12	416	LIPSEPALDSLVDPRVRSRKQPFVIYPVYDTAI
			\			DTKIHFSLLDGNVGEPDMSAGFCPNHKAAM
	}		ł			VLFLDRVYGIEVQDFLLHLLEGGFLPDLRAA
		l '	1			ASLDT/AEIGAMDFLLS*LFTLCLMMFFFIYPFI
	L	L :	L	L		NLLTMNVY
13	1363	A	249	535	105	WTFHRHLSPAPLIVCDQGTCVVSYYPQNIVQ
	}	}	<b>]</b> .			MPDTQMEQGLN/HLFLDGNA*PHSVECYCPS
		1	İ			TFEIAIKITSFVLYFHRYRAPEVLLRSSVYSSPI
		•				DVWAVGSIMAELYMLRPLFPGTSEVDEIFKIC
14	1264	<u> </u>	264		201	QVLGTPKKVSTLVPKLL
14	1364	A	254	572	201	YLLTXIGNLMMLLVINADSCLRTXM*FFLGH
			1		ļ	FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT   *SFFFIFSWGTNGCLLSAITYACYAAICHPLLS
			[			TMVMNRPLCTATVNATNKMGFLNSQVN
15	1365	Ā	257	425	68	THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM
1.5	1303	^	<i>""</i>	760	. 00	AIEFLLECDONIT\KLICENT*KNIAKNI*KRRV
[		7	1			TFTPIET*HPVKQMIKWQ*LTAWLRNRGYKKI
					i .	KOTPNSETAPSVCRNLVFDKCG
16	1366	A	263	104	481	FCIFRITEEDRGGDDCVVSVWTKORNNSCVK
	}	``			1	SKDVFSKPVNIFWALEESVLGVKARQPKPFFA
1	l	} '	1	1		AGNTFEMTCKYSSKNIKSPRYSVLIMAEKPV
		ļ				GDLSSPNETKYIISLDQDSVVKLENWTDASRV
17	1367	A	298	68	208	RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW
		l	·			VSSLAMKEMLTKTTM
18	1368	A	300	904	1	LVVGITGTRHHARVIFIFLVETGFPHVGQAGL
	L	L_			L	ELLTSGDPPALASQSAGITGMSHCARPKGHFG
						<del></del>

SEQ ID NO: of NO: of nucl- eotide seq- uence uence	a C-Custoina
nuclectide sequence uence uence uence in nucleotide location corresponding to last amino acid residue of residue of sequence u	cid
eotide sequence USSN 09/496 corresponding to last amino acid residue of peptide residue of sequence uence USSN 09/496 corresponding to last amino acid residue of peptide residue of sequence Y=Tyrosine, X=Unknown, *=Sto	=Histidine
seq- uence 09/496 correspondi to last amino M=Methionine, N=Asparagine, R quence 914 ng to first amino acid residue of peptide residue of sequence y=Tyrosine, X=Unknown, *=Str	
uence 914 ng to first acid residue Q=Ghutamine, R=Arginine, S=Sc amino acid of peptide T=Threonine, V=Valine, W=Try residue of sequence Y=Tyrosine, X=Unknown, *=Str	=Proline.
amino acid of peptide T=Threonine, V=Valine, W=Try residue of sequence Y=Tyrosine, X=Unknown, *=Str	
residue of sequence Y=Tyrosine, X=Unknown, *=Sto	
	on codon.
peptide /=possible nucleotide deletion, =	
sequence nucleotide insertion	
IHLK*MFYTMSQKMP*PTINL	ILLLIIPGNLNIF
KPNMGWLGPKTAFV*KDEV	
WK*DY*C/LQEVTDPIMEKGE	
GQPHQSTNALLRRCVR*RYH	LS\TVETAGLP*
KNTGHIPGQPFLFKLVFKC*N	VICI**QYKW*Q
NIGVKNKSFCPH*SSSPSL*FIG	GHHSRNP/CSFK
TEPHSVVQAGGQWRNLSSLQ	APPPGLMPLSR
ISLMSSWDYRRPPQ	
19   1369   A   302   3   445   NSPSRWAKIQMFEHTFCG*GG	CG/ER/NVHIHCS
WICRLRPLLWRAVREYLSKL	KNAELSFDPGV
KSRIWAVIQ/CIHLWDWLRKI	
AV*NKPRHLLSHIWKDVQNII	
20 1370 A 304 1 1339 FFFCGKEVPLFEQNKHPGPRA	
LLSAGEFTAGVGLSP*AIHSF	
GGPCHQPGGSPGPWMHTTQA	
GSSTWHQVPGQLGGSWGPRI	
CPHPPGFRLWMSPNQKPPTEI	
LMPGESPLIWEAEGKEDHLSI	
PLHSSLGNTVKP*PKNQKPKQ	
MAGQGQSRPAAR*PPCPALT	
RICRTVPGGPCPSPSGFRSCRI	
DAEPPSTPDTAPRCCTQSDTS	
CRALPGRICSAPAAGIRRARI PASPA A AS A B CRESWICEPOTA A	
PASPAAASARCPSWGPSCPAI AAPSRCTAWLRGEREPGPRPI	
VSFAPEVLSLPAVRQTKSWR	
ALVRSRGG	WIGHLEDITIG W
21 1371 A 326 799 1587 GSQVLPPPPSQDSATLPQDA*	GPRAAPGOPVC
E*GLQGAGVRRLRGEVLCQP	
HLSFSPROGAAPDTEPSAWGI	
LRHVRLFŠAGAPRGAATPCPI	
ARPMFRGHPPVRPLGPWGKV	
GVPAVQGECATKPSG*GL*PA	HLRGPPGPEVL
QWHWQLSAGRDPVPAEDPPI	
AAQAEPGADPEPEDKDQAAE	SRPAGAMSLSA
QGSGPVGGQGLR	1
22 1372 A 327 146 652 PHLENPHPEHSFPGAPLT*STL	SWSILSPREPSP
GAPCYPGHPHLENPHLEHLL1	
PGAPCYPEHPHLEHPLTWSTP	
SCRTPTRSILHRDHPLP*CLST	
	CPGRDSCYSVP
APPSTPLVLDVAPPGPQPASSO	
GTVVSP	
	IDDVVVCADNI I
GTVVSP	
23 1373 A 348 397 2 CIVSSCQGTRKPCHLEDANKI LQESL*VKQ*LIVAEKYVQILI NNEKRKMKKRKEEKKKCREI	rmqrrskwrr
23 1373 A 348 397 2 CIVSSCQGTRKPCHLEDANKI LQESL*VKQ*LIVAEKYVQILI NNEKRKMKKRKEEKKKCREI EEKKE*RREE\EERKKEKEDRI	rmqrrskwrr
23 1373 A 348 397 2 CIVSSCQGTRKPCHLEDANKI LQESL*VKQ*LIVAEKYVQILI NNEKRKMKKRKEEKKKCREI EEKKE*RREE\EERKKEKEDRI SRRLLRD	RMQRRSKWRR KERRKETSPRG
GTVVSP	RMQRRSKWRR KERRKETSPRG DALAPLDDSMP
GTVVSP	RMQRRSKWRR KERRKETSPRG DALAPLDDSMP ILLVSRVRGPQ
GTVVSP	RMQRRSKWRR KERRKETSPRG DALAPLDDSMP ILLVSRVRGPQ KRTENPERDQH
GTVVSP	RMQRRSKWRR KERRKETSPRG DALAPLDDSMP ILLVSRVRGPQ KRTENPERDQH
GTVVSP	RMQRRSKWRR KERRKETSPRG DALAPLDDSMP ILLVSRVRGPQ KRTENPERDQH WHWTNWIFTD
GTVVSP	RMQRRSKWRR KERRKETSPRG DALAPLDDSMP ILLVSRVRGPQ KRTENPERDQH WHWTNWIFTD RKSDEINQRTK
GTVVSP	RMQRRSKWRR KERRKETSPRG DALAPLDDSMP ILLVSRVRGPQ KRTENPERDQH WHWTNWIFTD RKSDEINQRTK
GTVVSP	RMQRRSKWRR KERRKETSPRG  DALAPLDDSMP ILLVSRVRGPQ KRTENPERDQH WHWINWIFTD  RKSDEINQRTK PNLIYKFNTISI
GTVVSP	RMQRRSKWRR KERRKETSPRG  DALAPLDDSMP ILLVSRVRGPQ KRTENPERDQH WHWINWIFTD  PRKSDEINQRTK PNLIYKFNTISI  DEQLEIEMNK
GTVVSP	RMQRRSKWRR KERRKETSPRG  DALAPLDDSMP ILLVSRVRGPQ KRTENPERDQH WHWINWIFTD  FRKSDEINQRTK PNLIYKFNTISI  NDEQLEIEMNK IIQNRHAENYKI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
28	1378	A	408	14	427	TICTNKFNNLDEIK/FLERHKLSKLTQEEVENL ITLKTSRETELVINK*VIPHKEKPGPDSFTGEF YQTFKEEL/II/ILHKLFQTIKYGRILPNSVYETSI TLKPKPEKDL\KENYRPLPLSNIDAK\LNKTLA NRI**HIR
29	1379	A	434	395	128	IYSKMCMERQRLNN*ILKKNKVRGIAVPDVK VYYKPTVIK/TSWIL*KDSHIVEWNRLENLEID PN/IKRLILDKGAEATEWRKDSFFRQWQ
30	1380	A	455	2	228	FFFETESHSVTQAGVQWCNPGFKRFSCFGLSS SWDYRYAPPRP\ANF\*FLVETGFYYVAQAGL KLLSPGDLPALAS
31	1381	A	462	393	2	QLMFDKGVKNIH\WGWTPPFTK*YWKNWISI CRRMNLNPYLSRYIKINSR\KDLTVRPEPIKLV EENTGKTIQDTGLGK*FIAKTSKAQSTKTNK* KRQTRYIKLK\KKSTASKENNRVKRQPLE*EK IFAN
32	1382	A	474	125	471	VKPYEIAVFLVKPIEYK*HILLSDPAIPLSGI*LK EIKAYT/RRICTPMFAAPVSVIA/RN*KQSK/CQ KQ*YVHRMEYYTTIKRSEILICTTTWVDFRNT ILRETDRIHKTTYDVISLI
33	1383	A	488	1825	2	KSACSFICSEEQPASPSPLKPGTYASETRPRDP HAAGPRRDSSEAETRRPRGA/DGSGTVVKGT PGSPAPPCSWGHGGVETEGAG*CPAAPGTDLR APGGSAGS*GLPSAGGSRGRKGWRAAGRQP STR*GRPGRHGGRGE*AGHPEPRQSALQSAG L/ASSPEPMGAALAEDGSGDSRGAGPRPQE*P PSVLSRS\GS*G*G*G*ASSPRSHSSRLGPP SAGFHGLRCGQPPFAAAPPGPWPGTGRPAGG AGSPPAAAGTAPPATRGAQSRRQNRTAGRNA SPQTAAGAGSPVQWALSRATG*TGETGSWC AGGTHQATHLTAAWVCPPTWSVRPGGSGPA AGLGR*GRHPAQSPPLPVPRG*PAWPQEAPSP SPASSEVALSSGSCWPDQAPGPARGSPPAPLA PAWPAAGRGRQR*GRQSAHPPPRR*STAVSL SGTS*WRRSP*AGTRTQQC*SPWLVPACSSRP L*RGTRRPSTQQSPQTTGTPGRSAGPGHPRS* GGRSPAGTGHLGAQTVASPH*GHWPTALSCL WASASPPGPEAPPQTGACIGTNCRYRAASAR RSSVAPACA*GWQ*AGSPPAVLRGPP*RVRER GALTHRPRAPDE
34	1384	A	497	422	2	APGASVGRAQAAEG*RGGPTGRPPSALGVS/E AGRAGRAGEGRPVPPAYPLCKSAQTSGPPKA RLS\PPLASCGGRGPPGGAACATCAPPAGPAR SSRCRRRSPPE*GPR*PSRPARPSPGSAASRRQ KLTPCRCQFRGLCA
35	1385	A	509	156	475	PTPYPGE*QAAFLLRGPGLRPPA/DPSLR/HRN LTELVVAVTDENIVGLFAALLAERRVLLTAS KLSTLTSCDHAFCALLYPMRWEHVLIPTLPPH LLDYC*CPPLPRT
36	1386	A	. 512	3	1631	FFFSFVCHLYCVSPTPGPHGRLATWL/PGLLA FLGLAAGGQTLCPAGELPGHARAQASGAPGS VLIAVPGRRRVHTCGPGPAAPSTRGECPPPAL GHTRPARPRPVVPFAPAVPQEPGGQGHGAA/P PATGHSAPRGCPPARAAPTGSATPAPPPAACA AFHSAWSVPPAGRQQG*RVPAPAFRRTTPGT PGQHLLDRPGAPPAQGSGPAPAPPPRLAGPA GPAAPPPGPPAASWHSSLSKSSSSLIGWSPPLP VGPGSLQ*TPPPQGPHLSGSCGGTSSWRGQR AAVARRLRSWNACGLSRVAGRSSASYPGRE

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:		nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	nou	in NO:	beginning nucleotide	location	
eotide			USSN	location		F=Phenylalanine, O=Glycine, H=Histidine,
1	seq-	l I			corresponding	Helsoleucine, Kelysine, Leleucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l .	,	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ĺ		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
}			ł	peptide	•	/=possible nucleotide deletion, \=possible
L	<u></u>			sequence	l	nucleotide insertion
	1					GRPSQSQ*PAGPPGMRGCCLRGW*PSSSGSD
			,	ļ	1	GPGPHPASTWLRAGKTGPSPPACGCA*LPPPS
						VSAAPQSPRTRCPRGCAAAAGLCVLAAAGAS
1				·		HGA\GLPGVRVHTQRVHIH*GAG/GCQTPRPR
1 . 1				Ì	[	LRSLPVLGLPAPRCPVSAHPWHRRSGSSCHA
1						ARLVPRHPAPGCP**TG*\PLITGFPEP*A*GLP
	ļ .			1	}	NHOAVGLEASGALQAGHRDELPTMVOLLDH
						SPDYPLKGRPHAP
37	1387	A	620	828	1	FRLPLAAGA/RGAAEPRVAVSMAPDPSAKIH
"	1307	^^	020	020	i •	WEASPEMQSKCHQKGKNNQTECFNHVRFLQ
1						RLNSTHLYACGTHAFQPLCAAIDAEAFTLPTS
1						FEEGKEKCPYDPARGFTGLIIDGGLYTATRYE
1						
	,					FRSIPDIRRSRHPHSLRTEETPMHWLNG*EDE
]					]	AQDDGG*GTISSFLLPWPADHPTPKSPGEPVH
[ - [						SIPVCCQVRGQPQSGGKESPACLKSLSNCLTH \DAEFVFSVLVRESKASAVGDDDKVYYFFTE
1 !						
20	1200		d20		400	RATEKESGSPTQSRSSHRVARGIPPL
38	1388	A	739	1	427	FRAMVSSTLKLGISILNGGNAEVQ/QGNRGKG
1						TSEEGKEG*EVPV*LPVSPPLPRPLQKMLDYL
]					ļ	KDKKEVGFFQSIQALMQTC\GEKVMADDEFT
1					•	QDLFRFLQLLCEGHNNDFQNYLRTQTGNTTT
<u></u>						INIIICTVDYLLRLQESI
39	1389	A	767	1	1030	TLDLTGPLLLGGVPNVPKDFRGRNRQFGGCM
						RNLSVDGKNVDMAGFIANNGTREGCAARRN
1						FCDGRRRQNGGTCVNRWNMYLCECPLRFGG
l i						KNCEQGEWPASSIPPVTAAWEALLLDVPGTT
i						VRGLHIQVRQPLVVYAAFTVDSHRPLQETVL
1			- 30			RRAPAPASGVPSPSGVGWDR*AGPAEPSPSTP
						ATVIISVPWYLGLMFRTR\KEDSVLMEATSGG
1	100					PTSFRLQVTGAPCHQGTC*VGARGRDPMLSG
Į į			į			LRVTDGEWHHLLIELKNVKEDSEMKHLVTM
						TLDYGMDQVSWHLHLLWG+TLPPAQGKTGA
						SEDKVSVRRGFRGCMQVRGGCGGRGEACPS
1 1						QAAPRL
40	1390	A	801	69	399	IHKIIIHKEDLNKWKYILCSGMERLSTVMIPVV
1 1		i i				PQIIYKFNA*Q\VILKFTW*E*GAKITTLRKNKL
						RGLVLVPLSTC*VKYLLDKVLPHIKTYYEAR
						VNKSVVLVQVTIM
41	1391	A	835	7	195	SMLKERKVFQFPSCLFFQYITWLGPPYHVLFD
				-		SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT
42	1392	A	841	1	415	GSTHASGYDKTPDFILOVPVAVEGHIIHWIES
"			···	•		KASFGDECSHHAYLHDOFWSYWNSLKHRTW
						OGIGTVASNLSOL*TLNAPFPELLLFRSLARTG
1 1				·		FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL
						7
142	1202		OAE	250	<u></u>	KACFPINIVIL
43	1393	A	845	358	92	PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG
1 . 1						LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP
						QRPMLPPSHAGLARPPPPEPISVP
44	1394	A	853	452	1	LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK
] !					1	SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS
]. [						PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG
						QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR
1						PLTFSTRRNVDPEIPERFR
45	1395	A	894	379	162	GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG
1						QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT
)						WLSMSMGK
46	1396	A	900	1	366	TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL
'	1330	^	<b>,</b> ,	•		
"	1390		500	•	, •••	EIQKYMRT/DQ+CVTHDISLYIVTKLALIFLIPR VFLFHQLNIT++CLHFFIMTTFIAIPFSFLFLGR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Głycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
47	1397	A	944	162	2	D/KSLAMLPRLVSNSWPQVILPP QLQNLASRGCL*SQLLRRLRRENRLNPGGGG CSEIAP\CTPAWVTQRDFFRKKK
48	1398	A	963	216	308	HFTPDRIAIVKNTRDSHCWRGC*EEGAPARC
49	1399	A	967	466	1	PRKRESWWGERLP/PRGFPPAAEDAPAPGWK GRKHASRTARAHVFHPIRQSIRSPVRGRPGDP RAAHTRSAGTRLQCKASRGG*GKGPAPTR*E GGPGSAPAPLPASSGCSLFPDSSPWTPPPPAPG AAAAOP**TPRCPAALRAGAHIGRVGRPY
50	1400	A	973	45 .	421	EKCIQALDVFVFCYIDHSSHCLMSCD*E/DQA LNFMPLEMEPKMSKLAFGCQRSSTSDDDSGC ALEEYAWVPPGLRPEQIQLYFACLPEEKVPY VNSPGEKHRIKQLLYQLPPHDNEVRYCQSLSE E
51	1401	*	992	2095	194	IRIRHEAARSCLGCAAGHVPAPGLRLLPTVRG PPGRRGPAAPGCVCY*SGESTFVSHVPQRMA WPGSAPPRGFHPLQSQTSPSDTVSSPQLSKEE DGPGWEHPLSSSL*SLGQAGGNH*QPEELAG WEPRGPPSLAPSSPT/TMWTALVLIWIFSLSLS ESHAASNDPRNFVPNKMWKGLVKRNASVET VDNKTSEDVTMAAASPVTLTKGTSAAHLNS MEVTTEDTSRTDVSEPATSGVAADGVTSIAPT AVASSTTAASITTAASSMTVASSAPTTAASST TVASIAPTTAASSMTAASSTPMTLALPAPTST STGRTPSTTATGHPSLSTALAQVPKSSALPRT ATLATLATRAQTVATTANTSSPMSTRPSPSKH MPSDTAASPVPPMRPQAQGPISQVSVDQPVV NTINKSTPMPSNTTPEPAPTPTVVTTTKAQAR EPTASPVPVPHTSPIPEMEAMSPTTQPSPMPYT QRAAGPGTSQAPEQVETEATPGTDSTGPTPRS SGGTKMPATDSCQPSTQGQYMV/DHH*APHP GRGRQNSPSGGAVTRGDPFHHSLGFVCPAGL *ELQEGLHPGGLLNQRDVCGLRNVRGAGA WREAWPLPRFFLLPLRPNQVLPNSFGAIEEIC QMLKHI
52	1402	A	994	1	462	ESGEFLVSFTLKKPTNVFHHINGMKPFNK/LIF *SHTDIAFYKIQHPFMLKALTKWA*EGT*PDR RYLH*SLRLNGEQLKTFPLRSGMR*G/CAILPL VLNAMLSIVPAVVPAGKTRHEKEITCPLIGQE EK*FS*FVGDMNTCVENKKESKKLLE
53	1403	<b>A</b>	1011	1	630	PEVIQOSAYDSKADIWSLGITAIELAKGEPPNS DMHPMRVLFLIPKNNPPTHCWRRLLESFKEV *LMLA*TKDPSIRPTAKELLKHKFIVKNSKKT SYLTELIDRFKRWKAEGHSDDESDSEGSDSES TSRENNTHPEWSFTTVRKKPDPKKVQNGAEQ DLVQTLSCLSMIITPAFAELKQQDENNASRNQ AIEELEKSIAVAEAAGPG
54	1404	A	1016	1	222	ISIDA*KAFDKIQH/CFMTTTLKKLGIDGKYLN TIKAIDDRHTVSTILNVEKLKAFL*RSGTRQRF PISGSGARI
55	1405	A	1033	3	366	HASVDGDEGSDDVYYYYTPAILRELQALNTA EAAEHRPEEDRMLSEDPWRPAHMIKGYMPL HNIPHTEVIDVTGLNQSHLYQHLNKGTPMKT OKRAA/LYTWHVLEQLEILRQINQQSHGPG
56	1406	A	1044	5	429	SVLTLOTRSPSKPLS\RKLMDWEVVSRNSISE DRLETQSRASRSPPVTPNQSQETPVDGKPLAL PPNQSQKNIRYHIHYLHLQYYLDRHISATLPIP SSSGIPTPIAVITDALTDLVELILGQPCSEESGR APGTLFLLAL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
F	uence	ŀ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uonce		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence	ļ <sup>*</sup>		914	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ľ	Į.	ł		residue of		Y=Tyrosine, X=Unknown, *=Stop codon,
		1	l		sequence	
	ł	ł	1	peptide	•	/-possible nucleotide deletion, \-possible
	1405	<u> </u>	40.50	sequence		nucleotide insertion
57	1407	A	1050	11	430	GAYAFETNGFPIMLVLTTDKIEGDVGIAGLYD
	Ì	1	j			MH\ISLPMAFLLRTLVRCTSYIIPVTHVLSTPV
ļ						TCLRRREKDGVIVDVLSDTASNHNGFPVEEH
	l	ł				ADDTHPARLQGPTLRSQPMGPLKHKAFEERA
						NLGLVQRRLRLED
58	1408	A	1058	258	419	LKHRDTPVVGANNRALSCTPLTSLTLCALCPL
	1 .	L				PCLGCPTXATCRLYQTTVAVVF
59	1409	Α	1064	3	425	KAFSFTTSLIGHQRMHTGERPYKCKECGKTF
		Ì				KGSSSLNNHQRIHTGEKPYKCNECGRAFSQC
		ļ			·	SSLIQHHRIHTGEKPYECTQCGKAFTSISRLSR
	i	Ì	Ì			HHRIHTGEKPFHCNECGKVFSYHSALIIHQRIH
	l .	[.		İ	i	TGEKPYACKDVGK
60	1410	A	1065	204	419	GGPPGPFLAHTHAGLQAPGPLLAPAGDEGDL
	-		1			LLLAVQQSCLADHLLTASWGGK/DPIPTKALG
	•	1	ĺ			EGOEGLPLTV
61	1411	A	1079	3	383	RHSRAHLCOPFHLVMRDLLQLGODIPQGCHY
	<del></del>	l	-3.7	-		LEENHLIHRDIAARNCLLSCAAPTRAATIGDF
l	1					GMARYIYRTRYYQLGDRAL/LPRKWMPPEAL
ļ	1	ł	1			LEGIFTYNTDSWTFGVLLWEIFSLGYMPYPGR
i						TN
62	1412	A	1080	1	859	VVEFLWSRRPSGSSDPRPRRPASKCOMMEER
02	1412	^	1000	1	637	ANLMHMMKLSIKVLLQSALSLGRSLDADHA
[						PLOOFFVVMEHCLKHGLKVKKSFIGONKSFF
1	Ī	}				GPLELVEKLCPEASDIATSVRNLPELKTAVGR
	i					1
1		1				GRAWLYLALMQKKLADYLKVLIDNKHLLSE
		İ				FYEPEALMMEEEGMVIVGLLVGLNVLDANIA
[ .	ſ	1				CLKGEDLDSQVGVIDFSLYLKDVQDLDGGKE
1	!	j				HERITOVLDQKNYVEELNRHLSCTVGDLQTK
		1	<u> </u>			IDGLEKTNSKLQERVSAATDRICSLQEEQQQL
	1413	<del> </del>	1083		615	REQUELIR
63	1413	A	1083	2	913	SSFAKHKRIHTGEKPFICLECGKAFTSSTTLTK
· ·	1	i				HRRIHTGEKPYTCEECGKAFRQSAILYVHRRI
[						HTGEKPYTCGECGKTFRQSANLYAHKKIHTG
		l				EKPYTCGDCGKTFRQSANLYAHKKIHTG\EKP
1	1	Ì	ľ			YKCKECGKAFKSYYSILKHKRTHTRGMSYEG
ļ	1	I	· .			DEC/QRSLN/RSSILSNHKIIHNEEK/PLKCEKCE
<u> </u>		<u> </u>				KAFNHTSICCRHKKN
64	1414	A	1084	946	1 .	KKQDLSSSLTDDSKNAQAPLALTESHLATLA
1	Į .					SSSQSPEAIKQLLDSGLPSLLVRSLASFCFSHIS
!	ł	I	ŀ			SSESIAQSIDISQDKLRRHHVPQQCNKMPITAD
ļ		ŀ		1		LVAPILRFLTEVGNSHIMKDWLGGSEVNPLW
l	Ì	ŀ	}			TALLFLLCHSGSTSGS\HNLG\AQQDQCKISFS
1		l				FFSWLTTGLTTQQRTAIE\NATVAFF\LQCI\SC
1	· ·	l	1			HPNNQKLMAQVLCELFQTSPQRGNLPTSGNI
}	l	İ				S\GFIR\RLFLQLMLEDEKVTMFLQSPCPLYKG
l		1 .				RINATSHVIQHP\MYGAGHKFRTLHLPVSTTL
	1					SDVLDRVSDTPSITAKLISKQKDDKKKK
65	1415	A	1087	103	324	PRAFEFVHTEMIVG/RVQNIHLFTLQVLEDRA
[		1			- <del>-</del> -	LFTMSVGSSLWSTYLIHVMALP/DRELLKPNA
<b>.</b>	<b>J</b>		j			SVALHKLSNALV
66	1416	A	1095	3	493	HETCSVTHIVSFSLPFLNPSHPASTPGHTENEQ
***	1 4410	l ^ _	1053	,	+73	PSLVWFDRGKFYLTFEGSSRGPSPLTMGAQD
1	}	i	l			
l	ł	ł				TLPVAAAFTETVNAYFKGADPSKCIVKITGE
		<b>l</b>				MVLSFPAGITRHFANNPSPAALTFRVINFSRLE
<u>}</u>	Ì	[· ·	ſ			HVLPNPQLLCCDNTQNDANTK\EFWVNMPNL
70	14.5	<b> </b>	1000			MTHLK
67	1417	A	1098	57	356	LKLTSLGFIIGVSVVGNLLISILLVKDKTLHRA
}		l		]		PYYFLLDLCCSDILRSAICFPFVFNSVKNGST
L	L	L	<u> </u>			WTYGTLTCKVIAFLGVLSCFHTAFMLFCISVT

SEQ ID   Not of Not of Not of nucleotide cotide sequence   Not of nucleotide cotide sequence   Not of nucleotide   Not of nu	SYDPKH OKPOPH SEGSFOF GAVRLG OPRPGSA APDVTAP EEPLRTL VRCVSR ALGSARAP NLSSVA SILKLDDD
ectide sequence ue	SYDPKH OKPOPH SEGSFOF GAVRLG OPRPGSA APDVTAP EEPLRTL VRCVSR ALGSARAP NLSSVA SLKLDDD YALFRNV
sequence uence	SYDPKH OKPOPH SEGSFOF GAVRLG OPRPGSA APDVTAP EEPLRTL VRCVSR LGSARAP LGSARAP LGSARAP
1419	SYDPKH OKPOPH SEGSFOF GAVRLG OPRPGSA APDVTAP EEPLRTL VRCVSR LGSARAP LGSARAP LGSARAP
amino acid residue of peptide sequence p	SYDPKH QKPQPH SEGSFQF GAVRLG QPRPGSA APDVTAP EEPLRTL VRCVSR LGSARAP NLSSVA SLKLDDD YALFRNV
residue of peptide sequence	SYDPKH QKPQPH SEGSFQF GAVRLG QPRPGSA APDVTAP EEPLRTL VRCVSR LGSARAP NLSSVA SLKLDDD YALFRNV
Peptide   Sequence	SYDPKH QKPQPH SEGSFQF GAVRLG QPRPGSA APDVTAP EEPLRTL VRCVSR LGSARAP NLSSVA SLKLDDD YALFRNV
Sequence	SYDPKH QKPQPH SEGSFQF JAVRLG QPRPGSA APDVTAP EEPLRTL VRCVSR LGSARAP NLSSVA SLKLDDD YALFRNV
RYL	QKPQPH SEGSFQF GAVRLG QPRPGSA APDVTAP EEPLRTL VRCVSR VRCVSR ALGSARAP NLSSVA SLKLDDD YALFRNV
1418	QKPQPH SEGSFQF GAVRLG QPRPGSA APDVTAP EEPLRTL VRCVSR VRCVSR ALGSARAP NLSSVA SLKLDDD YALFRNV
YEREGMQDWKTASGQSEEATQQSS   YTTYQSSSFLKYSSESHLLAWRENS    PGRSRARPPRTRQQRRGAAAGPGR    HPQSAAQPQLRAAARIPESPAAFPAG    RNSDASGPASLSRTLGRASSPRPPO/   SPAALAPRAARGGSRAAALAGAEA    APAPTRAAAPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP	QKPQPH SEGSFQF GAVRLG QPRPGSA APDVTAP EEPLRTL VRCVSR VRCVSR ALGSARAP NLSSVA SLKLDDD YALFRNV
YTTYQSSSFLKYSSESHLLAWRENS PGRSRARPPRTRQQRRGAAAGPGRK HPQSAAQPQLRAAARIPESPAAFPA RNSDASGPASLSRTLGRASSPRPPQ SPAALAPRAARGOSRAAALAGAEAI APRPTRAAAPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP	SEGSFOF GAVRLG OPRPGSA IPDVTAP EEPLRTL VRCVSR LGSARAP NLSSVA SLKLDDD YALFRNV
PGRSRARPPRTRQQRRGAAAGPGRC HPQSAAQPQLRAAARIPESPAAFPAG RNSDASGPASLSRTLGRASSPRPPQA SPAALAPRAARGGSRAAALAGABAA APRPTRAAAPPPPPPPPPPPPPPPPPPPPPPPPPPPP	GAVRLG OPRPGSA APDVTAP EEPLRTL VRCVSR LGSARAP NLSSVA SLKLDDD YALFRNV
HPQSAAQPQLRAAARIPESPAAFPA( RNSDASGPASLSRTLGRASSPRPPQ/ SPAALAPRAARGGSRAAALAGAEA APRPTRAAAPPPPPPPPPPPPPPPPPPPPPPPPPPPPP	OPRPGSA APDVTAP EEPLRTL VRCVSR LGSARAP NLSSVA SLKLDDD YALFRNV
RNSDASGPASLSRTLGRASSPRPPQ/ SPAALAPRAARGGSRAAALAGAEA APRPTRAAAPPPPPPPPPPPPPPPPPPPPPPPPPPPPP	APDVTAP EEPLRTL VRCVSR LGSARAP NLSSVA SLKLDDD YALFRNV
SPAALAPRAARGGSRAAALAGAEA APRPTRAAAPPPPPPPPLPPGAPPPP RARAPPWR/PAATGPPPRPVAPSRK APALQIRKGTSSGLPGRGGGSGPGN GNWRGSSFAVERPGMAKYQGEVQE SVIEGVSDQVLVAVVVSFALIATLV' HQNIHPENQELVRVLREQLQTEQDA QFYTDMYCPICLHQASFPVETNCGH PNSIW  69 1419 A 1107 2 466 FDTARLHEFGTSITQIFAVDNREDLQ FWQHFFDLSQWKHCCEELMKIEIMS LTKEATSVYHIDMSIDSPMKLESLTD TNGQFLIGGRESSLP/SS/CGPHSLMV RKRY/SYPASEPLHDEKGKKRQAPL TO 1420 A 1111 698 23 ALRRLHYVRATKVVFLSFRPFWREI SNTDRPSRMIFYPPREGALLLASYT AFAGLSREALRLALDDVAALHGPY DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHG' KSALRAAIKINSRKGFASDTASFEGI- QGHVHGVASSPSHDLAKEEGSHPPV QNTTHTRTSH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKAREAP EGYHSSTEGKAPAQQLPAFEELLAPI 72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	EEPLRTL VRCVSR LGSARAP NLSSVA SLKLDDD YALFRNV
APRPTRAAAPPPPPPPPPPPPPRAPPPP RARAPPWR/PAATGPPPRPVAPSRK APALQIRKGTSSGLPGRGGGSGPGM GNWRGSSFAVERPGMAKYQGEVQE SVIEGVSDQVLVAVVVSFALIATLV' HQNIHPENQELVRVLREQLQTEQDA QFYTDMYCPICLHQASFPVETNCGH PNSIW  69 1419 A 1107 2 466 FDTARLHEFGTSITQIFAVDNREDLQ FWQHFFDLSQWKHCCEEIMKIEIMS LTKEATSVYHDMSIDSPMKLESLTD TNGQFLIGQREESLP/SS/CGPHSLMV RKRY/SYPASEPLHDEKGKKRQAPLI TO 1420 A 1111 698 23 ALRRLHYVRATKVVFLSFRRPFWREI SNTDRPSRMIFYPPREGALLLASYT AFAGLSREEALRLALDDVAALHGPV DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHG KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPV QNTTHTRTSH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPPEQAGLSQFHLEPETQNPETITE QEAAAQLPQLPEVVELSSTKA\EAPA EGVHSSTEQKAPAQQLPAFEEILAPI 72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	VRCVSR LGSARAP NLSSVA SLKLDDD YALFRNV
RARAPPWR/PAATGPPPRPVAPSRK APALQIRKGTSSGLPGRGGSGPGN GNWRGSSFAVERPGMAKYQGEVQE SVIEGVSDQVLVAVVVSFALIATLV HQNIHPENQELVRVLREQLQTEQDA QFYTDMYCPICLHQASFPVETNCGH PNSIW  69 1419 A 1107 2 466 FDTARLHEFGTSITQIFAVDNREDLQ FWQHFFDLSQWKHCCEELMKIEIMS LTKEATSVYHDMSIDSPMKLESLTD TNGQFLIGQREESLP/SS/CGPHSLMV RKRY/SYPASEPLHDEKGKKRQAPL 70 1420 A 1111 698 23 ALRRLHYVRATKVFLSFRRPFWREI SNTDRPSRMIFYPPPREGALLLASYT AFAGLSREEALRALALDDVAALHGP DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHG KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPV QNTTHTRTSH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKAREAPA EGVHSSTEQKAPAQQLPAFEEILAPI 72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	LGSARAP NLSSVA SLKLDDD YALFRNV
APALQIRKGTSSGLPGRGGGSGPGN GNWRGSSFAVERPGMAKYQGEVQS SVIEGVSDQVLVAVVVSFALIATLV' HQNIHPENQELVRVLREQLQTEQDA QFYTDMYCPICLHQASIPVETNCGH PNSIW  69 1419 A 1107 2 466 FDTARLHEFGTSITQIFAVDNREDLQ FWQHFFDLSQWKHCCEELMKIEIMS LTKEATSVYHDMSIDSPMKLESLTD TNGQFLIGQREESLP/SS/CGPHSLMV RKRY/SYPASEPLHDEKGKKRQAPL 70 1420 A 1111 698 23 ALRRLHYVRATKVFLSFRRFFWREI SNTDRPSRMIFYPPREGALLLASYT AFAGLSREEALRILALDDVAALHGPV DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHGY KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPV QNTTHTRTSH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKANEAPA EGYHSSTEQKAPAQQLPAFEEILAPI 72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	NLSSVA SLKLDDD YALFRNV
GNWRGSSFAVERPGMAKYQGEVQE   SVIEGVSDQVLVAVVVSFALIATLV   HQNIHPENQELVRVLREQLQTEQDA   QFYTDMYCPICLHQASFPVETNCGH   PNSIW	LKLDDD YALFRNV
SVIEGVSDQVLVAVVVSFALIATLV	YALFRNV
HQNIHPENQELVRVLREQLQTEQDA QFYTDMYCPICLHQASFPVETNCGH PNSIW  69 1419 A 1107 2 466 FDTARLHEFGTSITQIFAVDNREDLQ FWQHFFDLSQWKHCCEELMKIEIMS LTKEATSVYHDMSIDSPMKLESLTD TNGQFLIGQREESLP/SS/CGPHSLMV RKRY/SYPASEPLHDEKGKKRQAPL 70 1420 A 1111 698 23 ALRRLHYVRATKVVFLSFRPFWREI SNTDRPSRMIFYPPPREGALLLASYT AFAGLSREEALRLALDDVAALHGPV DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHGY KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPV QNTTHTRTSH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKARAPA EGVHSSTEQKAPAQQLPAFEEILAPL 72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	
QFYTDMYCPICLHQASFPVETNCGH PNSIW  69 1419 A 1107 2 466 FDTARLHEFGTSITQIFAVDNREDLQ FWQHFFDLSQWKHCCEELMKIEIMS LTKEATSVYHDMSIDSPMKLESLTD TNGQFLIGQREESLP/SS/CGPHSLMV RKRY/SYPASEPLHDEKGKKRQAPL 70 1420 A 1111 698 23 ALRRLHYVRATKVVFLSFRRFWREI SNTDRPSRMIFYPPPREGALLLASYT AFAGLSREEALRLALDDVAALHGPV DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHG' KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPV QNTTHTRTSH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKARAPA EGVHSSTEQKAPAQQLPAFEEILAPL 72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKIL	TAAIKQ
69 1419 A 1107 2 466 FDTARLHEFGTSITQIFAVDNREDLQ FWQHFFDLSQWKHCCEELMKIEIMS LTKEATSVYHDMSIDSPMKLESLTD TNGQFLIGQREESLP/SS/CGPHSLMN RKRY/SYPASEPLHDEKGKKRQAPL RKRY/SYPASEPLHDEKGKKRQAPL ALRRLHYVRATKVFLSFRPFWREI SNTDRPSRMIFYPPPREGALLLASYT AFAGLSREEALRLALDDVAALHGPM DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHGM KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPM QNITHTRTSH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKANEAPA EGVHSSTEQKAPAQQLPAFEEILAPL  72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	T DOCUMENT
69 1419 A 1107 2 466 FDTARLHEFGTSITQIFAVDNREDLQ FWQHFFDLSQWKHCCEELMKIEIMS LTKEATSVYHDMSIDSPMKLESLTD TNGQFLIGQREESLP/SS/CGPHSLMV RKRY/SYPASEPLHDEKGKKRQAPL 70 1420 A 1111 698 23 ALRRLHYVRATKVFLSFRRPFWREI SNTDRPSRMIFYPPPREGALLLASYT AFAGLSREEALRLALDDVAALHGPM DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHQM KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPV QNTTHTRTSH 71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKANEAPA EGVHSSTEQKAPAQQLPAFEEILAPL 72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	LLCG2L1
FWQHFFDLSQWKHCCEEIMKIEIMS LTKEATSVYHDMSIDSPMKLESLTD TNGQFLIGQREESLP/SS/CGPHSLMV RKRY/SYPASEPLHDEKGKKRQAPL  70 1420 A 1111 698 23 ALRRLHYVRATKVFLSFRPFWRE SNTDRPSRMIFYPPPREGALLLASYT AFAGLSREEALRLALDDVAALHGPV DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHG KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPV QNITHTRISH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKAIEAPA EGVHSSTEQKAPAQQLPAFEEILAPL  72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	VUNAEA
LTKEATSVYHDMSIDSPMKLESLTD TNGQFLIGQREESLP/SS/CGPHSLMV RKRY/SYPASEPLHDEKGKKRQAPL  70 1420 A 1111 698 23 ALRRLHYVRATKVFLSFRRFFWREI SNTDRPSRMIFYPPPREGALLLASYT AFAGLSREEALRLALDDVAALHGPV DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHG' KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPV QNTTHTRTSH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKAIEAPI EGVHSSTEQKAPAQQLPAFEEILAPI  72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	
TNGQFLIGQREESLP/SS/CGPHSLMV RKRY/SYPASEPLHDEKGKKRQAPL  TO 1420 A 1111 698 23 ALRRLHYVRATKVFLSFRRPFWREI SNTDRPSRMIFYPPPREGALLLASYT AFAGLSREEALRLALDDVAALHGPM DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHGM KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPM QNITHTRISH  TI 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKAMEAPA EGVHSSTEQKAPAQQLPAFEEILAPL  T2 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	
RKRY/SYPASEPLHDEKGKKRQAPL  70 1420 A 1111 698 23 ALRRLHYVRATKV\FLSFRPFWREI SNTDRPSRMIFYPPPREGALLLASYT AFAGLSREEALRLALDDVAALHGPV DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHG' KSALRAAIKINSRKGPASDTASPEGH- QGHVHGVASSPSHDLAKEEGSHPPV QNITHTRTSH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKA\EAPA EGVHSSTEQKAPAQQLPAFEEILAPL  72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	
70 1420 A 1111 698 23 ALRRLHYVRATKV\FLSFRPFWREI SNTDRPSRMIFYPPREGALLLASYT AFAGLSREEALRLALDDVAALHGPV DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHG' KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPV QNITHTRTSH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKA\EAPA EGVHSSTEQKAPAQQLPAFEEILAPL  72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	
SNTDRPSRMIFYPPPREGALLLASYT AFAGLSREEALRLALDDVAALHGPV DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHGY KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPV QNITHTRISH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKARAPA EGVHSSTEQKAPAQQLPAFEEILAPL  72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	
AFAGLSREEALRLALDDVAALHGPY DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHG' KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPV QNITHTRTSH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKARAPA EGVHSSTEQKAPAQQLPAFEEILAPL  72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	
DGTGVVKRWAEDQHSQGGFVVQPI EKDDWTVPYGRIYFAGEHTAYPHG KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPV QNTTHTRTSH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKAVEAPA EGVHSSTEQKAPAQQLPAFEEILAPL  72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	
EKDDWTVPYGRIYFAGEHTAYPHO' KSALRAAIKINSRKGPASDTASPEGH QGHVHGVASSPSHDLAKEEGSHPPV QNTTHITRISH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKAIEAPA EGVHSSTEQKAPAQQLPAFEEILAPL  72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	
QGHVHGVASSPSHDLAKEEGSHPPV QNTTHTRTSH  71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKA\EAPA EGVHSSTEQKAPAQQLPAFEEILAPL  72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	
71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKA\EAPA EGVHSSTEQKAPAQQLPAFEEILAPI 72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	ASDMEG
71 1421 A 1119 2 385 QKQTLQNGYLDSSMDILYLGSLPPE PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKA\EAPA EGVHSSTEQKAPAQQLPAFEEILAPL 72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	'QGQLSL
PPGPPEQAGLSQFHLEPETQNPETTE QEAAAQLPQLPEVVELSSTKA\EAPA EGVHSSTEQKAPAQQLPAFEEILAPL 72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	
QEAAAQLPQLPEVVELSSTKA\EAPA EGVHSSTEQKAPAQQLPAFEEILAPL 72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	
EGVHSSTEQKAPAQQLPAFEEILAPL   T2   1422   A   1127   1   906   HAQYVGPYRLEKTLGKGQTGLVKL	
72 1422 A 1127 1 906 HAQYVGPYRLEKTLGKGQTGLVKL	
	LIHHE
[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]	
EHPHVLKLHGVYHNKKYFPPDELTS QVSPHGKLSARRSWDLLSGFPRYLV	
GELFDYLVKKGRLTPKEARKFFRQI	
HSYSICHROLKPENLLLDEKNNIRLA	
LQVGDSLLETSCGSPHYACPEVIKGE	
RADMWSCGVILFALLVGALPFDDDD	
KVKRGVFHMPHFIPPDCQSLLRGMI	
LSLEQIQKHPWYLGGNFIS	
73 1423 A 1128 1 802 LRNALDVLHREVPRVLVNLVDFLNE	TIMROV
FLGNPDKCPVQQA/MLEPLGSKTETI	
MPITCPTONEPFLRTPRNSNYTYPIKI	
SDFLCTEWKASNSVPTSVHOLRPAD	
LGDSLTTAVGARPNNSSDLPTSWRG	UK V V A A
GDGNLETHTTLPNILKKFNPYLLGFS	
TAGLNVAAEGARARDMPAQAWDL	LSWSIG
SPDINLEKDWKLVTLFIGGNDLCHY	LSWSIG TSTWEG
HLATEYVQHIQQALDILSE	LSWSIG TSTWEG VERMKN
74 1424 A 1139 60 480 FREPCLLVPGDHQPLREASWLA/LPP	LSWSIG TSTWEG VERMKN CENPEA
DSPLCCVEVAIPCNKGAHSVGLKGW	LSWSIG TSTWEG VERMKN CENPEA
VLGMRDTIPQEHPWESTPDLCFCRD	LSWSIG TSTWEG VERMKN CENPEA
EQPAADAAVAKGEF/QGEQIAPVPA\	LSWSIG ITSTWEG VERMKN CENPEA IGLWGT ILLAQG
AADPAPVHITAHPKGA	LSWSIG ITSTWEG VERMKN CENPEA IGLWGT ILLAQG PEEIEVE
75 1425 A 1147 2 413 PFPHQHPQEP\KGSCWPQSALRGQCF	LSWSIG TSTWEG VERMKN CENPEA IGLWGT TLLAQG PEEIEVE IIAAHPE
TTTSDLCSLQVPVSSHRNPLLDLAAN	LSWSIG TISTWEG VERMKN CENPEA IGLWGT ILLAQG PEEIEVE IIAAHPE GPVLGV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  RFDNFSSLSIQWESTRPVLASIEPELPMQLVSQ DDESGQKKLHGLQAILVHEASGTTAITATAT
76	1426	A	1155	38	410	GYQESHLSSAR PIISAPAQDDPILLSFIHCLHANLLCVWRRDVK PDCKEIWIFWWGDEPNLVVQYIMNCMLWK KDSGKMAFPMNVGRC/FFKEIHNLLERCLMD KNFVLIGKWFVRPYYKDEKPVNKSEHLSCAF T
77	1427	A	1162	526	350	RFPQGLEDVSTYPVLIEELLSRGWSEEELQGV
78	1428	A	1171		1293	LRGNLLRVFRQVEKVQEENKWQSPLED MAESASPPSSSAAAPAAEPGVTTEQPGPRSPP SSPPGLEEPLDGADPHVPHPDLAPIAFFCLRQT TSPRNWCIKMVCNPWFECVSMLVILLNCVTL GMYQPCDDMDCLSDRCKILQVFDDFIFIFFA MEMVLKMVALGIFGKKCYLGDTWNRLDFFI VMAGMVEYSLDLQNINLSAIRTVRVLRPLKA INRVPSMRILVNLLLDTLFMLGNVLLLCFFVF FIFGIIGVQLWAGLLRNRCFLEENFTIQGDVAL PPYYQPEEDDEMPFICSLSGDNGIMGCHEIPP LKEQGRECCLSKDDVYDFGAERQDLNASGL CVNWNRYYNVCRTGSANPHKGAINFDNIGY AWIVIFQVITLEGWVEIMYYVMDAHSFYNFI YFILLIUSVREPGLLGGSFSTAQSPKCQGDSFP GVAAESLLLRGWVLWLPGGG
79	1429	A	1175	1	405	PNDFFKDMFPDLPGGPLGPIKAENDYGAYLN FLSATHLGGLFPPWPLVEERKLKPKASQQCPI CHKVIMGAGKLPRHMRTHTGEKPYMCTICE VRFTRQDKLKIHMRKHTGERPYLCIHCNAKF VHNYDLKNHMR
80	1430	A	1182	25	198	EMNELSQQLSQQGGRGASQCPSPPAPTLPNPT PLCQLQLQRVNTGLPTPPCHPGAGAA
81	1431	A	1186	254	583	KTVLDVGAGTGILSIFCAQAGARRVYAVEAS AIWQQAREVVRFNGLEDRVHVLPGPVETVEL PEQVDAIVSEWMGYGLLHESMLSSVLHARTK VVKDGGFFLPXSSELFM
82	1432	A	1187	2	716	DFVDAARNLPLESTKSPAEPSKSVPSLENDPRA SSQGLPSQGPVQNQGRRGEQRPKKF/TVIQHT SSFEKSDSLEQPSGLEGEDKPLAQFPSPPPAPH GRSAHSLQPKLVRQPNIQVPEILVTEEPDRPD TEPEPPPKEPEKTEEFQWPQGSQTLAQFPVEK LPPKKKRLGLAKMAQSSGESSFESSVPLFRSP SQESNVSLSGSSRSALFERDDHGKAEAPSPSF DMGPKPLGTHMLTV
83	1433	A	1188	517	804	ESPGLSKVLRTGAFAYPFLFDNLPLFYRLGLC WGRGHGCGQEALSTSHGYHLFCALLTGFLFA SHLPERLAPGRFDYIGHSHQLFHICAVLGTHF Q
84	1434	A	1192	45	476	LGDVGFWVERTPVHEAAQRGESLQLQQLIES GACVNQVTVDSTIPLHAASLQGQARCVQLLL AAGAQVDARNIDGSTPLCECLRLGQHRVCEA LAVLRGQGQPSPVHSVPPARGLHXREFRMC* GFLFDVGXNLEAHEFHFGEP
85	1435	A	1194	69	410	KRSEEASAPPFPLGGTGAAPTRASLPEQILLPR SCLEARKSQPDEKLLSALHNSRTWN*EPRRSQ HRLVSPEVHPGRRGSSPGVAECKLTSAYFRT GRSPCPSLPGTTRTNSLL
86	1436	A	1215	3	405	LPSHTCGNPGRLPNGIQQGSTFNLGDKVRYSC NLGPFLEGHAVLTCHAGSENSATWDFPLPSC RADDACGGTLRG/AEWHHLQPPLPLG/ATKN

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ĺ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ł	ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	İ	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
}	ļ		ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		peptide		/-possible nucleotide deletion,possible
[		1		sequence		nucleotide insertion
					·	NADCTWTILAELGDTIALVFIDFQLEDGYDFL
ľ		ĺ		İ	1	EVTGTEGSSLW
87	1437	A	1216	226	964	GTARFGPMVGFGANRRAGRLPSLVLGVLLV
		l		ļ		VIVVLAFNYWSISSRHVLLOEEVAELOGOVO
1 '	}				ļ	RTEVARGRLEKRNSDLFAVVGHAQETDRPEG
			1		Į	GRLRPPQQPAAGQRGPREEMEDDKVKLQNN
					]	ISYQMADIHHLKEQLAELRQEFLRQEDOLOD
		ĺ	[		ĺ	YRKNNTYLVKRLEYESFQCGQQMKELRAOH
1	1					EENIKKLADQFLEEQKQETQKIQSNDGKELDI
					ł	NNQVVPKNIPKVAENVADKNEEPSSNHIPHG
88	1438	Α	1218	1	534	PEFGTTISCGYLMATDVSRRPSVHKAVEIEOE
1						RVKSAGAWIIHPYSDFRFYWDLIMLLLMVGN
1						LIVLPVGITFFKEENSP\PWIVFNVLSDTFFLLD
					·	LVLNFRTGIVVEEGAEILLAPRAIRTRYLRTW
1			ĺ	*		FLVDLISSIPVDYIFLVVELEPRLDAEVYKTAR
						ALRIVRFTKILSLLRL
89	1439	Α	1223	1	743	MGFDEVFMINLRRRQDRRERMLRALQAQEIE
						CRLVEAVDGKVGMLTRSNAAPGRHLAMLET
1 1						LVVVAPRFVDADNLILNPDTLSLLIAENKTVV
						APMLDSRAAYSNFWCGMTSQGYYKRTPAYI
]						PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNL
J J						VAFYPPHPDYTWSFDDIIVFAFSCKQAEVQMY
						VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ
90	1440	A	1227	2	349	LEVMVPSSPSSAQSMAVVSADHIGLVISYL
30	1-1-10	^	122/		349	NKTSFIFYLKNIVVADLIMTLTFPFRIVHDAGF
					i	GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/
1 1						YDRY/WKVVRHL/WDSWMTGI/SFTRVYLLG LGARLVWFGKLILAKGGHGGISWL
91	1441	$\overline{\mathbf{A}}$	1245	3	1937	LGSSDVRAPQRSELGAESPSRMVASQAYNLT
					1557	SALTPILTRSRVLNEEPLTLAGF\SRAPANLSD
<b>j</b> ]		J	1	}		VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF
1 1		1				QTQAGAQIPIERLASERATTVKVPNNSDWAAR
						GHRSSANSV\VQPQAFVGAVVTLDSSNPAAV
i l	Ì		i		i	LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR
i i	ľ	- 1	ľ			PNEHNCSASRRIRPESLQGADHRPYTFFISPGT
1 1	ı İ	1	ļ			RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC
	ŀ	1	ł			QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR
		ŀ		j		HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML
	j	į	1		•	TCAVCLVTYMVMAAILHKLDQLDASRGRAIP
	ſ	ĺ	1	Ĭ	ſ	FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM
		ſ	Į		i	LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP
	F	l				HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD
		ł	İ		ł	LQTARSTFFLVNDWLSVETEANGGLVEKEVL
· 1		l				AASKASFRVPTPS\AALLRFRRLLVAELQRGF
	J	}	-			FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL
1		· ]		ĺ	j	FLGANAVWYGAVGDSAYSTGRVSRLNPLSV
		İ	]	]		DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV
, <b>!</b>	Į			Ì		GWGWGPGSTGNGAWASAPCPEPPLSSAAAR
<u> </u>						GKGVHQRLLGKGQHT
92	1442	A	1246	5	562	VFDEENILNELNDPLREEIVNFNCRKLVATMP
. I		- 1	Į.	1	ļ	LFANADPNFVTAMLSKLRFEVFQPGDYIREG
. 1	- 1	l	- 1	1	1	AVGKKMYFIQHGVAGVITKSSKEMKLTDGS
	Į		ĺ	j	l	YFGEICLLTKGRRTASVRADTYCRLYSLSVD
	ĺ	ļ	İ			NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN
<del></del>			10.0	100		SILLQKFQKDLNTGVFNNQENEILKQIVKH
93	1 4 4 7		1740	180	901	TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL
	1443	A	1249	100		
	1443	^	1249			PGRKASCSTAGSGSRGLPPASSPMVSSAHNPN
	1443	A	1249			

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LAPPSGERSRLARGSTIRSTFHGGQVRDRRAG
						GGGGGVQNGPPASPTLAHEAAPLPAGRPRP TTNLFTKLTSKLTRRVADEPERIGGPEVTRRP RQEDHLSPGGRGCSEL
94	1444	A	1261	3	385	KFSQWGLTKPKLSNASP/WISLVKKLMKKWS VTQNLTFREQLEAGIRYFDLRVSSKPGDADQ EIYFIHGLFGIKVWDGLMEIDSFLTQHPQEIIFL DFNHFYAMDETHHKCLVLRIQEAFGNKLCPA CR
95	1445	A	1282	2	550	GPRDNPG\EDPRFEIVEHFGIAWFTFELVARFA VAPDFLKFFKNALNLIDLMSIVPFYITLVVNL VVESTPTLANLGRVAQVLRLMRIFRILKLARH STGLRSLGATLKYSYKEVGLLLLYLSVGISIFS VVAYTIEKEENEGLATIPACWWWATVSMTT VGYGDVVPGTTAGKLTASACILA
96	1446	A	1294	1	1456	QLLPPSNRENAGLLVGRCLCSAALRPVGDLIT SSGQVAVRNAPQAGSAKAGKGKFQDNFEFIQ YFKKFFDANCNEKDYNPVAAGQGQETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA ABTTYMIF/LV/LVILLGSLYLVTLILAV/VAMA YEEQNQATLEEABQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRRKKRKQKEQSGGEEKDED EFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMKKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE
97	1447	A	1295	2	2057	IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDHRESR WMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESBTSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTILEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR
98	1448	A	1304	118	453	SGPSSRAIYLHRKEYSONLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS
99	1449	A	1306	3	1660	CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE TGTGTAYEGFLSVPRPSGVRRGWQRVFAALS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
	! ! !			amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=-possible nucleotide insertion
				ocquence		DSRLLLFDAPDLRLSPPSGALLQVLDLRDPQF SATPVLASDVIHAQSRDLPRIFRVTTSQLAVPP TTCTVLLLAESEGERERWLQVLGELQRLLLD ARPRPRPVYTLKEAYDNGLPLLPHTLCAAILD QDRLALGTEEGLFVIHLRSNDIFQVGECRRVQ QLTLSPSAGLLVVLCGRGPSVRLFALAELENI EVEVPKIPESRGCQVLAAGSILQARTPVLCVA VKRQVLCYQLGPGPGPWQRRIRELQAPATVQ SLGLLGDRLCVGAAGGFALYPLLNEAAPLAL GAGLVPEELPPSRGGLGEALGAVELSLSEFLL LFTTAGIYVDGAGRKSRGHELLWPAAPMGW
	·					GYAAPYLTVFSENSIDVFDVRRAEWVQTVPL KK\VRPLNPEGSLFLYGTEKVRLTYLRNQLAE KDEFDIPDLTDNSRRQLFRTKSKRRFFFRVSE EQQKQQRREMI.KDFFVRSKLISPPTNFNHLV HVGPANGRPGARDKSP
100	1450	A	1318	918	190	SLCVPGPVDTGTFAVMSVMVGSVTESLAPQA LNDSMINETARDAARVQVASTLSVLVGLFQV GLGLIHFGFVVTYLSEPLVRGYTTAAAVQVF VSQLKYVFGLHLSSHSGPLSLIYTVLEVCWKL PQSKVGTVVTAAVAGVVLVVVKLLNDKLQQ QLPMPIPGELLTLIGATGISYGMGLKHRFEAG\ PPVAPNTQLFSKLVGSAFTIAVVGFAIAISLGK IFALRHGYRVDSNQVWVMRDV
101	1451	A	1353	220	445	DWPDLFTYPLIGSPKCFQSARPEVRMYRRTVR SSHGNHALQEVLPRSGHGTEFTKQKHLEAAD HGHPPARMSIFSR
102	1452	A	1363	542	2	AHLLMLNLAL\TDLL\YLTSLPFLIHYYASGEN WIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRY CVIIHPMSCFSIHKTRCAVVACAVVWIISLVA VIPMTFLITSTNRTNRSACLDLTSSDELNTIKW YNLILTA\LLCLPLVIVTLCYTTIIHTLTHGHAN \DSCLKQKARRLTILLL
103	1453	A	1371	2	410	CHSTESSSDFILPGDYLLGGLCPLHSGCLQVC SFNEHGYHLFQAMRLAVEEINNSTALLPNITL GYQLYDVCSDSANVYATLRVLSLPGQHHEL QGDLLHYSPTVLAVIGPDSTNRAATTAALLSP FLVPMLLEQ
104	1454	A	1376	3	432	NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG KSREERFCNENTPCPVPIF
105	1455	A	1379	2	396	GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD IILAVIDSIFVWFIFISLAQIMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA
106	1456	A	1383	1	432	EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP
107	1457	A	1386	719	558	FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL
108	1458	A	1397	631	2	RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK AYVCSILASQLLIIPVFAIGGANVWVLGLLLF

NO: of   NO: of   No: of   Not   n	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine O=Cysteine,
Docation   Docation							
Sequence   1949   1	nucl-	peptide	Ì		nucleotide	location	
10			]				
	, -	uence					
residue of peptide sequence	uence			914			Q=Glutamine, R=Arginine, S=Scrine,
		1					
				ĺ		sequence	r=1 yrosine, X=Unknown, =Stop codon,
109						1	micleotide insertion
		<del> </del>	<b></b>			<del></del>	
1459	ľ		!	ļ	}	Ì	
1459			<u> </u>	i			LSFSLTFVVILRNRRPGKSLVR
RYTSKKSAKYRLQGTIPRGDVSLTIMESESDS	109	1459	A	1402	15	387	VLVALPDT\VTSETVVTEVLGHRVTLPCLYSS
110							WSHNSNSMCWGKDQCPYSGCKEALIRTDGM
110				ľ	Í	f	
LAELAFPYGVLATCA*SLLSC*YCVILFPCSCF   FFHSPDALFSLLLSCYFSYCFFYYLFFSSEP   CLLLASSPPLFILLASL   CLLLASSPPLFILLASL   CLLLASSPPLFILLASL   CLLLASSPPLFILLASL   CLLLASSPPLFILLASL   CLLLASSPPLFILLASL   CLLLASSPPLFILLASCA   CLAKPLD SYLINSSSSSTATAGGGIGGS   CSSNPVATFVFQQSSDPVSSYGFVNTAESST   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   SPSLLFSQDSKLATTS	110	1460		1401		250	
	110	1400	A	1421	3	350	HEDLSSLLTRGSGNQERERQLKKLISLRDWM
CLLLASSPPFLILLASL							FEUGDDALEGIALI GOVEDOVOREVVI EEGGODE
111			•	i			
QCALKPDLSYLNNSSSSSTPATSAGGGIFGSS   TSSNPPVATPYFGQSSDPVSSYGPVNTAEST   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   STSLFSQDSKLATTS   STSLFSQDSKLATTS   STSLSSSSSSSSCGGGGGFGFTTWWR SRRSSQRTCSRAGGAGWSRTWFRSS*TSSSC   STSCSSSSSSSSCGCGGGGFGFPLGARGVHITSCLINSC   MSSTTSTTSTF   HEDIMTHYDRLVDE*ALNAGKQRYEKMISG   MSSTTSSTTSTF   HEDIMTHYDRLVDE*ALNAGKQRYEKMISG   MSSTTSSTTSTF   HEDIMTHYDRLVOVLFYVSFYLFQSCDINVL   GIFLTFLLSNFLIVCVLLFYVSFYLFQSCDINVL   GIFLTFLLSNFLIVCVLLFYVSFYLFQSCDINVL   GRIKVHLPGHKTGPAVAKDTPEPVKKEFTVP   ATSQGP*SFFSEEPPLPPSEQPEPVRYGQDLINLO   QRTKVHLPGHKTGPAVAKDTPEPVKKEFTVP   ATSQGP*SFFSEEPPLPPSEQPEPVRYFLPP*EPQS   EPP*KNA*LKQMHAATTHWQQHQVQTQC   QYHGIMQ   QYHGIMQ   GRKCGQYWPLEKDSRRFGFLTVSNLGVEN   MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN   NYCN   MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN   NYCN   MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN   NYCN   MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN   MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN   MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN   MYCN   MNHYKWGDLSFTRAAGTAGGLLYSNLQ   HLKWNGDSLFLCLSLPC   HLKWNGDSLFLCLSLPC   GREGGLLYSNLQ   HLKWNGDSLFLCLSLPC   GREGGLLYSNLQ   HLKWNGDSLFLCLSLPC   GREGGLLYSNLQ   HLKWNGDSLFLCLSLPC   GREGGLLYSNLQ   HLKWNGDSLFLCLSLPC   GREGGLLYSNLQ   HLKWNGDSLFLCLSLPC   GREGGLLYSNLQ   HLKWNGDSLFLCLSLPC   GREGGLLYSNLQ   HLKWNGDSLFLCLSLPC   GREGGLLYSNLQ   HLKWNGDSLFLCLSLPC   GREGGLASVPPVCMPLLNTTQLAKLEPPNS GREGGLASVPPVCMPLLYTERPPWQNPLPVNRGQAR   GREGGLASVPPVCMPLLYTERPPWQNPLPVNRGQAR   GREGGLASVPPVCMPLLYTERPPCMPLT   GREDP*KPPPCMPL   HLKWL*GNPFYPCMPLT   GREDP*KPPPCMPL   GREDP*KPPPCMPL   HLKWL*GNPFYPCMPL   GREDP*KPPPCMPL   HLKWL*GNPFYPCMPL   GREDP*KPPPCMPL   GREDP*KPPPCMPL   HLKWL*GNPFYPCMPL   GREDP*KPPCMPL   GREDP*KPPPCMPL   GREDP*KPPPCMPL   GREDP*KPPPCMPL   GREDP*	111	1461	A	1426	2	344	
TSSSNPPVATFVFGQSSPVSSYGFVNTAESST   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   SDSLLFSQDSKLATTS   STSSUFFGGFGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1	•			<u> </u>		QCALKPDLSYLNNSSSSSSTPATSAGGGIFGSS
112	{·						TSSSNPPVATFVFGQSSDPVSSYGFVNTAESST
SRRSSQRTCSRACSGAWSRTW*RSS*TSSSSC   STSCSSSSSRSCGRPGGPLGARGVHITSCLNSC					<u> </u>		SDSLLFSQDSKLATTS
STSCSSSSRSCGPGGPLGARGVHITSCLNSC   MSSSTTSSTTSTT   MEDIMTHYDRI.VDE*ALNAGKQRYEKMISG   MYLGEIVRNILDFTKKGFLLRGQISEMLKTR   GIFLTFLISNFLIVCVLLFYVSFYLFQSCINFVL   GGPLTFLISNFLIVCVLLFYVSFYLFQSCINFVL   QRTKVHLPGHKTGPAVAKDTPEPVKKEFTVP   ATSQGP*SFSEEPPPSNEVPPTLPP*SEQS   EDP*KNA*LKQMHAATTHWQQHQQHQVGC   QYHGIMQ   QYHGYYDEXBTGGFLTVSNLGVEN   MNHYKKSTLEILNPEVNPGFFLTLWKQGEN   NYCN   QHWKSCHWGVVQKRRAL*VYSFEEG   GRRKCGGVWPLEKDSRIFGFLTVSNLGVEN   QHWLEGHLCSMSLQEFTRAAGTAGQLLYSNLQ   HLKWNGGSLFLCISLPC   QHWLEGHLCSMSLQEFTRAAGTAGQLLYSNLQ   QLQATSVPIHPVCMPLNNTQKSKQPLPSAPEN   QLQATSVPIHPVCMPLNNTQKSKQPLPSAPEN   NPEEELASDPNNESL*RPWALEDFEIGRPLG   KGK   QLQATSVPIHPVCMPLNNTQKSKQPLPSAPEN   NPEEELASDPNNESL*RPWALEDFEIGRPLG   QHPWLEGHTCLDNNIHQAASEPINNNFAESKR   NLAFLATGVVRHMRKLFMGANLEGPGPTVS   H     1469   A 1486   1 398   GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL   NINVSYPPTKQLTYBEQDLGWKFRYYLTNOE   KALTKFLKWVNWDLPQEAKQALELLGKWK   PMDVKDSLELLSSHYTNPTVRRYAVARLRQA   DDEDLLMYL   DDEDLMYL   DDEDLMYL   DDEDLMYL   DDEDLMYL   DDEDLMYL   DDEDLMYL   DDEDLMYL   DDEDLMYL   D	112	1462	Α	1434	46	372	
MSSSTTSSTTSTF	] [						
113	]						
MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR   GIFLTFLLSNFLLVCVLLFYVSFYLFQSCINFVL	113	1463	_	1420	2	202	
114	11.5	1405	A	1437	3	292	
114	1 1						GIFL TELL SNEL IVCVLL EVVSEVLENSCHMENT
QRTKVHLPGHKTGPAVAKDTPEPVKKEFTVP	114	1464	A	1463	1	396	KOOAVPEPHSSTTTPOEOFONWYGODLLNIO
ATSQGP*SPFSEEPPLPPSNEEVPPTLPP*EPQS EDP*KNA*LKQMHAATTHWQHQQHQVGC QYHGIMQ  115 1465 A 1464 291 2 AGSYPSMVWSCHWGVTQKRRAL*VYSFEEG GRRKCGQYWPLEKDSRIRFGFLTVSNLGVEN MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN NYCN  116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ YWTKYQVWEWLQHFLDTNQLDANCIPFQEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC  117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY							QRTKVHLPGHKTGPAVAKDTPEPVKKEFTVP
115	1						ATSQGP*SPFSEEPPLPPSNEEVPPTLPP*EPQS
115 1465 A 1464 291 2 AGSYPSMVWSCHWGVTQKRRAL*VYSFEEG GRRKCGQYWPLEKDSRIRFGFLTVSNLGVEN MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN NYCN  116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ YWTKYQVWBWLQHFLDTNQLDANCIPFQEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC  117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEELASDPNNESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY	l i						
GRRKCGQYWPLEKDSRIRFGFLTVSNLGVEN MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN NYCN  116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ YWTKYQVWBWLQHFLDTNQLDANCIPFQEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC  117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVFLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFCDQFL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY	115	1466		1424	001		
MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN NYCN	113	1465	A	1464	291	2	AGSYPSMVWSCHWGVTQKRRAL*VYSFEEG
116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ YWTKYQVWBWLQHFLDTNQLDANCIPFQEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC  117 1467 A 1479 1 381 GTSGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY							
116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ YWTKYQVWEWLQHFLDTNQLDANCIPFQEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC  117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY	ĺ						
YWTKYQVWEWLQHFLDTNQLDANCIPFQEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC  117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY	116	1466	A	1465	667	337	
DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC  117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY		ľ					YWTKYOVWEWLOHFLDTNOLDANCIPFOEF
117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY						l	DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ
VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY							
QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY	117	1467	A	1479	1	381	
NPEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY	l [						
Italia	ŀ			- 1		ŀ	
118 1468 A 1485 3 385 TYLWL*GNPFFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY			j				
PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 I 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY	118	1468	$\overline{\mathbf{A}}$	1485	3	385	
QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 I 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY					-		
NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 I 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY			ŀ				
119 1469 A 1486 I 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY			J	-			
NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY							
KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY	119	1469	Α	1486	1	398	GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL
PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY	ſ	J	1	ĺ			
DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY	1	ļ	1	. [			
120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY		1	Į	1			
industrial of the second of th	120	1470	<del>  </del>	1407	<del></del>	2000	
, , , , , , , , , , , , , , , , , , ,	120	14/0	^	147/	3	צעע	
LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL	ĺ	1	i		ĺ	ì	
RTSPLYDRLDAQGARWMEKHGFERPKYFVP							
PDKDLLALEQSKTFYKPDWFDIVESEVKCCK	}		· }	ļ		Ì	
EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS		į	- 1	ľ		• 1	EAVCVIDMSSFTEFEITSTGDQALEVLOYLFS
	, ,		i			1	NDLDVPVGHIVHTGMLNEGGGYENDCSIARL
NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	1	J	ł	ł	1	ł	
LLLEDVTWKYTALNLIGPRAVDVLSELSYAP		l					LLLEDVTWKYTALNLIGPRAVDVLSELSYAP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	[	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	vence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	"		914	ng to first	acid residue	Q=Ghtamine, R=Arginine, S=Serine,
		1.	7.1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	}	1		peptide	Sequence	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
		-		Sequence	<del> </del>	MTPDHFPSLFCKEMSVGYANGIRVMSMTHT
						GEPGFMLYIPIEYRWGFTMLSTLVSNS
121	1471	A	1498	3	306	AQFLLVGWDHIL*LIVL*TNLTELGRTTCDQN
1	• • • • • • • • • • • • • • • • • • • •	1	1170	1	300	WPNSPDVLNHGCFYMQCLSKDCTIGYVSRE
				ľ	ļ	MLVAHTHTVEEHTGTHLQYVSWPDHSVPDD
1		ł	1	1	ł	SSDFVEFEN
122	1472	A	1533	121	329	LGLFSFVWTEVLEEPKDFSCETEDFKTLHCT
		•	1000	121	323	WDPGTDTALGWSKQPSQSYTLFES*VGSGYII
l		ŀ				DNFFLA
123	1473	A	1547	111	408	DARTTWKPRNGSSGIWPGDGAK*PPAVEQAE
		**	1347	***	406	RGHVEMIEKLTFLNLHTSEKDKGGNTALHLA
						AKHGHSPAVQVLLAQWQDINEMNEKQQTPL
						HVAADRG
124	1474	A	1555	1	745	MTFDDDDKNTYGVALVWKKFQTQSLRLSDL
1	14/4	n l	1333	1	/43	HRKSHLWRGIVSITLIEGRDLKAMDSNGLSDP
1 1						YVKFRLGHQKYKSKIMPKTLNPQWREQFDF
i i						HLYEERGGVIDITAWDKDAGKRDDFIGRCQV
ļ						DLSALSREQTHKLELQLEEGEGHLVLLVTLT
i i						ASATVSISDLSVNSLEDQKEREEILKRYSPLRI
] . ]				,		FHNLKDVGFLQVKVIRAEGLMAADVTGKSD
1 1						PFCVVELNNDRLLTHTVYKNLNPEWNKVFIL
						*VALVWKKFQTQSLRLSDLHRKSHLWRGIVS
1 1				•		ITLIEGRDLKAMDSNGLSDPYVKFRLGHQKY
1						KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT
1 1	ł					AWDKDAGKRDDFIGRCOVDLSALSREOTHK
			•			LELQLEEGEGHLVLLVTLTASATVSISDLSVN
						SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV
	1					KVIRAEGLMAADVTGKSDPFCVVELNNDRLL
	•	i	i			THTVYKNLNPEWNKVFTL
125	1475	A	1556	57	509	GGPAPNSRYAEP*KNSLAMT*AHADCENYVA
!!			1			CGGLDNICSIYNLKTREGNVRVSRELPGHTGY
1	j		,			LSCCRPLDDSQIVTSSGDTTCALWDIETAQOT
1	1	1			•	TTFTGHSGDVMSLSLSPDMRTFVSGACDASS
	1	- 1	- 1			KLWDIRDGMCRQSFTGHVSDINAVS
126	1476	A	1592	3	178	KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL
			[			EMLPTCDLADQHNIKFHYAFALNR*ER
127	1477	A	1612	1	497	TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS
[ ]		l		i		VLGAYISFGVPSSHLLTASVMSAPASLAAAKL
[		- 1	!		ļ	FWPETEKPKITLKNAMKMESGDSGNLL*AAT
[ [	ſ	ĺ	[	ſ	•	QGASSSISLVANIAVNLIAFLALLSFMNSALA
		l	į			WVGNMFDYPQLSFELICSYIFMPFSFMMGVE
[	l			1	}	WPDSFM
128	1478	A	1619	286	486	CCMNSKAQESVFKNVLCNPPALSEMPDVKA
, 1		ļ	- 1	ļ		EDEVDFRASSISEEVAVGSIAATLKMKQGPM
		ľ	1	i		TQAINR
129	1479	A	1627	1	395	PTRGALRYWIFGRFLCNIWAAVDVRCCTATI
l l	ļ	}		1	-	MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA
	i	i				LLCVWALSLVTYIGPLLGWRHPAPEDETICQI
	i	Ţ	-		l	NEEPGYVLFSTPGSFYLPLAIMLVMN*RVYRV
[		ĺ		ŀ	1	AKTE
130	1480	Ã	1638	2	466	DPRVRTKIVNRKTTIYEIODKTGSMAVVGKG
ŀ		l		ĺ		ECHNIPCEKGDKLRLFCFRLRKRENMSKLMS
		I			İ	EMHSFIQIQKNTNQRSHDSRSMALPOEOSOHP
·}	ļ	ļ			]	KPSEASTTLPESHLKTPOMPPTTPSSSSFTKVT
		. ]	1	l	1	KDKDIK*LLFNLYSSVEILPEVLHLKT
131	1481	A	1651	607	3	LAEGGDVFDCVLNGGPLPESRAKALFRQMVE
					- 1	AIRYCHGCGVAHRDLKCENALLQGFNLKLTD
	İ			[	ì	FGFAKVLPKSHRELSQTFCGSTAYAAPEVLQ
		- 1	i	ì	ł	GIPHDSKKGDVWSMGVVLYVMLCASLPFDD
			——			

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning		Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in NO.		nucleotide	D=Aspartic Acid, E=Glutamic Acid,
			I	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ľ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ì	j	ŀ	ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ı		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ł		l		peptide	•	/=possible nucleotide deletion, \=possible
Í		ĺ	Í	sequence	1	nucleotide insertion
				-		TDIPKMLWQQQKGVSFPTHLSISADCQDLLK
1	f	İ		ł	1	DITENDE WOODE VON THE SISAUCQUELK
1			l			RLLEPDMILRPSIEEVSWHPWLAST**KQWQV
120	1400		1222			LSNKVGGESKPKKKK
132	1482	A	1656	150	48	LVAKSLLYCGCLFFLLQLAKNVGNNSFNDIM
1 '						EANLTSPSPKPTPSSDM*VFLIY*TYFGAWHV
			I	Ĺ	ſ	VDAQ
133	1483	A	1660	3	406	RKHIKLLIQKLSDVP*ECQNNQL*KLTEICEKE
1			}	!	1	KKEFKKKMDDQRPEKITEA*SKDKSPMEEEK
1					Ì	TEMIRSYIQEVGRYIKRLEEAQSKRLEKLREK
j .						HKEIRQPILDEKPKGEGSSSFLSETCHEDTSWF
1						PNFTP
134	1484	Α	1666	1276	466	PGSTHASARITTY*L*IILSNATEVDNNFSKPPP
1	1-10-7	43	1000	12/0	1 700	
						FFPAGAPPASSSSSSSSSSSPPTVSTAPPLIPPPGF
1 .						PPPPGAPPPSLIPTIESGHSSGYDSRSARAFPYG
						NVAFPHLPGSAPSWPSLVDTSKQWDYYARSS
1						SSSSSSSSSSSSPRDRDRER*RTREREREDHS
<b>i</b> i				171		PTPSVFNSDEERYRYREYAERGYERHRASRE.
1						KEERHRERRHREKEETRHKSSRSNSRRRHESE
1						EGDSHRRHKHKKSKRSKEGKEAGSEPAPEQE
						STEATPAE
135	1485	Α	1673	1	417	PTRPVNSSQAFALVYYTLGALGGNLIAHMGL
1				_	127	GYRYWAGIGVLQSCESALTHYRLVANHVAS
						DISLTGGSVVQRIRLPDEVENPGMNSGMLQE
1	1					DLIQYYQFLAEKGDVQAQVGLGQLHLHGGR
136	1486	A	1678	525	9	GV*QNHQRAFDYFNLAA
130	1400	^	10/6	323	9	ANTSLSSAAVSAVSPPPCRTSTATTLPPPMPSF
1 1	1			ĺ		FCVFPSPSMSPSPSEFLSCIASVSRVHSLSSSSS
			1			GSSSTASSLNFSAIMGSSSATASWVLSTASTPP
!!!		ł		l l		CPSALPSSPAQES*SLAASSSAWPVAGISPSGA
l					•	CTFPAGSASGAAKAPSPSWRCPSFRALFSLLD
L				j		SSSLSL
137	1487	A	1680	1	2999	AHRDEIQRKFDALRNSCTVITDLEEQLNQLTE
1 1		1		•		DNAELNNONFYLSKOLDEASGANDEIVOLRS
	1	ĺ		1		EVDHLRREITEREMQLTSQKQTMEALKTTCT
			1			MLEEQVMDLEALNDELLEKERQWEAWRSVL
1	1	1		l l	l	GDEKSQFECRVRELQRMLDTEKQSRARADQ
ļ l	1	1				RITESRQVVELAVKEHKAEILALQQALKEQK
] ]	j	· ]	ļ	l		LKAESLSDKLNDLEKKHAMLEMNARSLQQK
1 1		ĺ	1	. [		I BEDDEL KUDI I BEUVAI UUU WI UNUUMAKSI QUK
1 · 1	ļ	ł		İ	•	LETERELKORLLEEQAKLQQQMDLQKNHIFR
( I	i	ľ	i	ľ	l	LTQGLQEALDRADLLKTERSDLEYQLENIQV
, [	1	I	ļ	l		LYSHEKYKMEGTISQQTKLIDFLQAKMDQPA
) <b>i</b>	- 1	j	į	ļ	J	KKKKVPLQYNELKLALEKEKARČAELEEALQ
[	1	1	- 1	• [		KTRIELRSAREEAAHRKATDHPHPSTPATARQ
; <b>j</b>	j	j	1	ŀ	•	QIAMSAIVRSPEHQPSAMSLLAPPSSRRKESST
) i		ŀ	į	ĺ	. [	PEEFSRRLKERMHHNIPHRFNVGLNMRATKC
j J	}		l	l	I	AVCLDTVHFGRQASKCLECQVMCHPKCSTC
1 1	i	- 1	ľ	f	ł	LPATCGLPAEYVTHFTEAFCRDKMNSPGLOT
} <b>!</b>	- 1		1	l	Į	KEPSSSLHLEGWMKVPRNNKRGQQGWDRK
, ,	1		j	ŀ	j	YIVLEGSKVLIYDNEAREAGQRPVEEFELCLP
[	ì	l	i	l		DGDVSIHGAVGASELANTAKADVPYILKMES
ı l	į	į	Ţ			HPHTTCWPGRTLYLLAPSFPDKQRWVTALES
) (	ļ	- 1	1	{	•	
l I	i	}	l		ļ	VVAGGRVSREKAEADAKLLGNSLLKLEGDD
1 1		- 1	- 1	j	. 1	RLDMNCTLPFSDQVVLVGTEEGLYALNVLK
1 }				1		NSLTHVPGIGAVFQIYIIKDLEKLLMIAGEERA
j.	Ì	- 1	1	İ	İ	LCLVDVKKVKQSLAQSHLPAQPDISPNIFEAV
l í	- (	ĺ	- [	i	l	KGCHLFGAGKIENGLCICAAMPSKVVILRYN
{	1		ļ			ENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNK
1 1	1	ł	1		ļ	FYEIDMKQYTLEEFLDKNDHSLAPAVFAASS
1 1		1			1	NSPPVSIVQVNSAGQREEYLLCFHEFGVFVDS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
138	1488	A	1686	2	526	NSLEVIEIQARSSAGTPARAYLDIPNPRYLGPA ISSGAIYLASSYQDKLRVICCKGNLVKESGTE HHRGPSTSRR*PASPLPQYQGQRAFLQGRRK GRPQGPAPGAGSPPESGPGLWAALGCSLVWV
						PLCCLGGAAGRL*ARSGKSGLRRRAHAGPP PGGPCNSCP*CSAPESGGRGPLPGPGTGGVCS CWTRGCQTTARTAAAAAAAPGPAGRRPPGGA PQNGSCAASASQEAAAPPPMCPPGRRWAVAS PPETRCPAAPGTRCRRLEAA
139	1489	A	1693	3	376	LPSMSNCTSCFRLQSRTES*IRQAGHLLGRNE FIETKALGCAWFSLCYYLVLYFESSHKVDFVF IV*CFSTPPGAQMTIMSQACAERCNIMRLVDR RWAGIAKOVGTQKIIGRVHLGEQKALGL
140	1490	A	1704	3	376	ERTNKFIKELIMDGKNLIAATKSLSVAQRKFA HSLRDFKFEFIGDAVTDDERCIDASLREFSNFL KNLEEQREIMVS*EGCKLISQLSRGKKIWIWK LVLVEVVKHLSLGTVVHCNGKMRFPEP
141	1491	A	1743	1	362	LITNKVFVARELSCLDVHLDSTGSTAVVADQ DKLELELVLKGSYEDTQTSFLGTASAFRFHY MAAL*TELSGRLRSSKSNGWNGDNSTGYLTV PLRPLTIVKEVTMDVPAPNVRGLNWMG
142	1492	A	1769	•	406	NNPSTLPRGS*PMSPRITMGRRRQRRREHKSS LSLASSTVOPGGQIVHTETTEVVLCGDPLSGF GLQLQGGIFATETLSSPPLVCFIEPDSPAERCG LLQVGDRVLSINGIATEDGTMEEANQLLRDA ALAHKVV
143	1493	A	1789	1	447	QMLRNGGDQNTVPDYHFADRIRELL*PTEDQ KNCIP*DTYLRPSALGNIVEEVTHPCSPGPCPA NELCEVNRKGCTSGDPCLPYFCVQGCKLGQA SDFIARQGTLIQVPSSAGEVECYKICSCGQSGL LENCMEMHCMDLPTDTSALVR
144	1494	A	1814	1	404	PGRRFRPRLSQAGTDSGS*VFPDSFPSAPAEPL PYFLQEPQDAYIVKNKPVELRCRAFPATQIYF KCNGEWVSQNDHVTQEGLDEATGLRVREVH IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK SRRAYVRI
145	1495	A	1827	26	448	XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THLALCPIVQHPEDTCIHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL
146	1496	A	1828	574	333	QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET
147	1497	<b>A</b>	1855	1	372	ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q
148	1498	A	1879	568	7	LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS
149	1499	A	1880	611	24	GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP VQVISGEFEARLIYQGVAHTTGGADQRLVVD

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in NO:	nucleotide	location	D=Aspertic Acid, E=Glutamic Acid, F=Phenylalanine. G=Glycine. H=Histidine.
cotide	sed-	}	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uciic		914	ng to first	acid residue	
Gence	ŀ	1.	714	amino acid	1.	Q=Glutamine, R=Arginine, S=Serine,
1		ľ		residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		1			sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	l			peptide		/-possible nucleotide deletion, \-possible
<b></b>	<del></del>	<del> </del>	<b></b> -	sequence		nucleotide insertion
i			ĺ			IGGASTELVTGTGAQTT*LFSLSMGCVTWLER
1	l		Ì			YFADRNLGQENFDAAQKAAREVLRPVADEL
1	}	1	1	9		RYHSWKEVRGASVTVQALQEIMMAQGMDE
150	1500	<u> </u>	1004		750	RITMEIWPVD
150	1500	Α	1894	2	750	GRVDFFHTDYRPLIRDSNNYVLDEQTQQAPH
						LMPPPFLVDVDGNPHPTKYQRLVPGRENSAD
	]	j	j			EHLIPQLGYVATSDGEVIEQIISLQTNDNDERS
	l					PESSILDGMIRQLQQQQDQRMGADQDTIPRG
1	1	ļ '	1			LSNGEETPRRGFRRLSLDIQSPPNIGLRRSGQV
1						EGVRQMHQNAPRSQIATERDLQAWKRRVVV
						PEVPLGIFRKLEDFRLEKGEEERNLYIIGRKRK
						TLQLSHKSDSVGLVSQSRPRTCRRKYP
151	1501	A	1900	141	785	GKTIQIQTIMQNKYKTVQKQYKTIPKNKKA
	İ	<b>!</b>				MEMQIKKQFQDTCKVQTKQYKALKNHQLEV
}	•					TPKNEHKTILKTLKDEQTRKLAILAEQYEQSI
						NEMMASQALRLDEAQEAECQALRLQLQQEM
1						ELLNAYQSKIKMQTEAQHERELQKLEQRVSL
			•	,		RRAHLEQKIEEELAALQKERSERIKNLLERQE
						REIETFDMESLRMGFGNLVTLDFPKEDYR
152	1502	A	1915	2	377	LVRLLDTQRDGLQNYEALLGLTNLSGRSDKL
J j						RQKIFKERALPDIENYMFENHDQLRQAATEC
						MCNMVLHKEVQERFLADGNDRLKLVVLLCG
J I					·	EDDDKVQNAAAGALAMLTAAHKKLCLKMT
						QVTT
153	1503	Α	1921	1	237	AYQSLRLEYLQIPPVSRAYTTACVLTSAAVQL
						ELITPFQLYFIPELIFKHFQIWRLITNFLFFVPFG
<u> </u>						FNFLLYMIFLYT
154	1504	A	1928	2	354	EMVEGGEGKMCINTEWGGFGDNGCIDDIRTR
1	,					YDTEVDEGSLNPGKQRYEKMTSGMYLGEIV
1 1	į					RQILIDLTKQGLLFRGQISERLRTRGIFETKFLS
<b></b>	1505		4000			QIESDRLALLQVRRILQQLGLD
155	1505	A	1929	2	369	TEIAKIKMEAKKKYEKELTMFQNDFEKACQA
1 1				'		KSEALVLREKSTLERIHKHQEIETKEIYAQRQ
						LLLKDMDLLRGREAELKQRVEAFESYQLELK
1-1	1506		1005		a÷a	DDYHRTYRLIEDDRINIQISGHWQESP
156	1506	A	1935	1	270	VTRKLPIFTVDAFTARAFRGSPAADCLLENEL
		]				DEDMHQKIAREMNLSETAFIRKLHPTDNFAQ
J.,	1505					RSCFGLIWFTPTTDLQILTSSILPSIL
157	1507	Α	1936	584	305	ESKVNNEKFRTKSPKPAESPQSATKQLDQPTA
]	·	[				AYEYYDAGNHWCKDCNTICGTMFDFFTHMH
- <u></u> -	1.505					NKKHTQGQFQKSSDFQKEELQQTFLPPERQG
158	1508	A	1939	1	423	TTHRLNVTAEPPCTSMPIYWMPDVPHRCTTA
	. ' [	' I	İ	l	i i	NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED
<b>[</b>						RLHCQTQACPPLSWPQRLDILLGTARAIQFLH
( l		1	Ì	i		QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA
لــــا		اـــــــــــــــــــــــــــــــــــــ				RFSRFAGSSPIQSSM
159	1509	A	1974	3	401	HTSTARLLLHRGAGKEAVTSDGYTALHLAAR
			ļ		[	NGHLATVKLLVEEKADVLARGPLNQTALHL
1		ł				AAAHGHSEVVEELVSADVIDLFDEQGLSALH
		)		ļ	ļ	LAAQGRHAQTVETLLRHGAHINLQSLKFQGG
لـــــا						HGPAATLLR
160	1510	A	1982	2	417	KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA
'		ŀ				IYSEYCNNHPGACLELANLMKQGKYRHFFEA
] . ]			1	İ		CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL
i (	ĺ		i	ľ		LKYTTQEHGDYSNIKAAYEAMKNVACLINER
			]			KRKLESIDKIA
161	1511	A	1984	4	770	RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE
[		·			1	LWLETLQGQRHSHTGVKSTPGQSAAILMKLR
<u> </u>	{	1	i	ł	4	SSHNASKTLNANNMETLIECQSEGDIKEHPLL

mucleotide   seq	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
USSN   Osadon   Osa	NO: of	NO: of	hod	ID NO:	beginning		D=Aspartic Acid, E=Glutamic Acid,
Sequence			l	1			
1914   ng to first amino acid or residue of peptide popular of peptide peptide popular of peptide pept			l				
amino acid residue of peptide squence   peptide squence   peptide   squence   peptide   squence   peptide   squence   peptide   squence   peptide   squence   peptide   squence   peptide   squence   peptide   squence   peptid		uence		1			M=Methionine, N=Asparagine, P=Proline,
residue of   peptide	uence		)	914			
peptide   sequence		1	l				
Sequence	j	]	]			sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ASCSSEDSICQLEVKRYKVISWPFLARRES   PASIFSGALETIUKASLEPOPLSICIOSDITI.PRIQUILTILCI.KGPSTEGIFRAANEKARKEI.  SPEINSGALVDIELISLEPOPLSICIOSDITI.PRIQUILTILCI.KGPSTEGIFRAANEKARKEI.  SPEINSGAVOLIELII.PVILLAVVEVEPULSIP   RKLLSSDLPEWMGALEMQDEEDRIBALK   RELINSGAVOLIELII.PVILLAVVEVEPULSIP   RKLLSSDLPEWMGALEMQDEEDRIBALK   RKLLSSDLPEWMGALEMQDEEDRIBALK   RKPQRWGTVCDDIMMIDHASVICKQLECGSA   VSPSGSSNPGEGSGPTWFDLLCKGNESALWN   CKHQQWGKHCDIABAGNGVCSSAU   CKHQQWGKHCDIABAGNGVCSSAU   CKHQQWGKHCDIABAGNGVCSSAU   CKHQQWGKHCDIABAGNGVCSSAU   CKHQQWGKHCDIABAGNGVCSSAU   CKHQQWGKHCDIABAGNGVCSSAU   CKHQQWGKHCDIABAGNGVCSSAU   CKHQQWGKHCDIABAGNGVCSSAU   CKHQQWGKHCDIABAGNGVCSSAU   CKHQQWGKHCDIABAGNGVCSSAU   CKHQQWGKHCDIABAGNGVCSSAU   CKHQQWGKHCDIABAGNGVCSAU   CKHQQWGKHCDIABAGNGVCSAU   CKHQQWGKHCDIABAGNGVCSAU   CKHQQWGKHCDIABAGNGVCSAU   CKHQQWGKHCDIABAGNGVCSAU   CKHQQWGKHCDIAGAGNG   CKHQQWGKHCDIAGAGNG   CHAPACH	ŀ	Į.					
PASDFSGALETDLKASIEPOPLSIGGDSDITG   PRIQUITITICLKGSFTEGIFRRANBEARKEL   REELINSGBAVDLERLPYHLLAVYPKDFTRSP					sequence		
			ĺ		( .	1	ASCESEDSICQLIEVKKRKKVLSWPFLMRRLS
			1	1			PASDFSGALETDLKASLFDQPLSIICGDSDTLP
		•	1	i	ł	ł	
1512   A   1986   864   501		1			,		
163	162	1512		1006	974	F01	
VSPSGSSNFGEGSGPWTDDLICNGRIESALWN   CKRIGGWOKHNCDHAEDAGVISSKD	102	1512	A	1986	804	201	
1513   A   2001   419   187		ĺ	1		ĺ		
1513   A   2001   419   187	ļ		1				
LASRSNIAFMGTI. VECGKAKGVVIGTGENSE	162	1512		2001	410	100	
FGDINLSTEVVIS	103	1513	A	2001	419	187	
1514		·	ĺ	·			· ·
SHIYKRDSFANKFIKIQAIEILKIRKPNDIETTKI	164	1514		2012	204		
ENNWYFVVADSSKAGFTTIYKWERETGFYSH	104	1514	A	2012	284	397	
1515   A   2013   2   403   EDPELGHFYDYPMALFSTFELFLTIIDGPANY   NVDLPFMYSITYAAFAIIATILMINILIAMMG   DITHWRVAHERDELWRAQIVATTYMLERKLP   RCL.WPRSGIGGREYGLGDRWILRVEDRQDLN   RQRIQRYA   RQRIQRYA   RQRIQRYA   CQQEGLGLKAVVQILLSHGRNGLPGEFASS   QGLSAASSTPVFHLALQIDSAPDNIDWVEMLF   NKNMYTERLQNYMVLEQCFSDSSSL.YRFLTY   SYLLAFNYWLLLAFVTLCYDWQVGSFL.VETI   WDMRNLATIFLAVVMALLSLHCLAFKRIE   HKEVLVGLLFL.YPFPHSNJFFRVGFVVARR   VLYMPSMGYCILFVHGLSKLCTWLNRCGATT   LIVSTVILLILLFSWKTVKQNEIWSRESLFRS   GVQTLPHNAKVHTNYNAPILKDQGRNKEAIY   HYRTALNNNKAWDYLCWRFRKTLTDLP   AASSASSLTVTIGRLSACSHSILRSEGFGA   ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW   PEVVLPDPVEBTRHHAEVVKKVNEMIVTGQY   GRIFAVVHFASRQWKVTSEDLLIIGNELDLA   CGERRLEEVILLVGADNFTLLGKDLGKDLV   RVBATVIERTESWRFMRFKKRNFKKKRIV   TTPQTVLRINSIEIAPCLL   TTPQTVLRINSIEIAPCLL   ROBERT   TTPQTVLRINSIEIAPCLL   ROBERT   TVSTUSGEGGPJ_GHNVFFFSSGRLGALFIRGI   EDNISRSKRGLFHENGLVKINNDLVONGVFS   KKQQAQTOTSEPPQDLOVEFFGQQDQVLR   RQGAARVFMPLQAQVKAKASKPLQMQIKA   PPRLRAARVLMPLQAQVRAFRLLQVOSQVS   KKQQAQTOTSEPPQDLOVEFFGQDQDVLR   TVSTUSSGEGGPJ_GHNVFFFSSLSSRLGLFIRGI   EDNISRSKRGLFHENGLVKINNDLVDKTFA   QAQDVFRQAMKSPSVLHVLPPQNREQYEKS   VIOSLNIFGNNDGVLKTKVPPPVHGKSGLKTA   NILTGTDSPFTDASSALQNKSFRVPRLGGKPS   SPSLSPLMGFGSNKNAKKIKDLKKGPFGGLFF   TVVTRDSSHGFGPFTVMFROSH   TVVTRDSSHGFGPFTVMFROSH   TVVTRDSSHGFFFTVMFROSH   TVVTRDSSHGFFFTVMFROSH   TVVTRDSSHGFFFTVMFROSH   TVVTRDSSHGFFFTVMFROSH   TVVTRDSSHGFFFTVMFROSH   TVVTRDSSHGFFFTVMFROSH   TVVTRDSSHGFFFTVMFROSH   TVVTRDSSHGFFFTVMFROSH   TVVTRDSSHGFFFTVMFROSH   TVVTRDSSHGFFFTVMFROSH   TVVTRDSSHGFFTPTCVMFROSH   TVVTRDSSHGFFTPTCVMFROSH   TVVTRDSSHGFFTPTCVMFROSH   TVVTRDSSHGFFTPTCVMFROSH   TVVTRDSSHGFFTPTCVMFROSH   TVVTRDSSHGFFTPTCVMFROSH   TVVTRDSSHGFFTPTCVMFROSH   TVVTRDSSHGFFTPTCVMFROSH   TVVTRDSSHGFFTPTCVMFROSH   TVVTRDSSHGFFTPTCVMFROSH   TVVTRDSSHGFTPTCVMFROSH   TVVTRDSSHGFTPTCVMFROSH   TVVTRDSSHGFTPTCVMFROSH   TVVTRDSSHGFTPTCVMFROSH   TVVTRDSSHGFTPTCVMFROSH   TVVTRDSSHGFTPTCVMFROSH   TVVTRDSSHGFTPTCVMFROSH   TVVTRDSSHGFTPTCVMFROSH   TVVTRDSSHGFTPTCVMFROSH   TVVTRDSSHGFTPTCVMFROSH   TVVTRDSSHGFTPT							
1515		Ċ	1				
NVDLPPMYSITYAAFAIIATILMINILIAMMO	166	1515	<u> </u>	2012		400	
DTHWRVAHERDELWRAQIVATTVMLERKLP	102	1515	I A	2013	Z ·	403	
RCL WPRSGICGREYGLGDR WILRVEDRQDLN							
RQRIQRYA							
1516						. 1	
QGLSAASSTPVFHLALQIDSAPDNIDWVEMLF NKNMVTERLQNVMVLEQCFSDSSLYRFLTY SYLLAPNVWLLLAPVTLCYDWQVGSIPLVETI WDMRNLATIFLAVVMALLSLHCLAAFKRLE HKEVLVGLLFLVFFFIPASNLFFRVGFVVAER VLYMPSMGYCLFVHGLSKLCTWLNRCGATT LLVSTVLLLLIFSWKTVKQNEIWLSRESLFRS GVQTLPHNAKVHYNYANFLKDQRRKEAIY FYRTALNNNKAMDYLCWEFRKTLTDLP  167 1517 A 2025 696 71 AAASAASSLTVTLGRLASACSHSILRPSGPGA ASLWSASRFINSQSTSYLPGYVPKTSLSSPPW PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY GRLFAVVHFASRQWKVTSEDLLIGNELDLA CGERIRLEKVLLVGADNFTLLGKDLV RVBATVIEKTESWPRIMRFRKRKNFKKKRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAQVADAPLLGVQSQVS KQQAQTQTSEPVDLDQVPEEFQQDQVLR PFLIRRAARVLMPLQAQVRAPRLLQVQSQVS KQQAQTQTSEPVDLDQVPEEFQQDQVLR PFLIRRAARVLMPLQAQVRAPRLLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLHVLPPQNREQYEKS VIGSLNFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSFLMGFGSNNAKKIKIDLKKGPEGLGF TVVTRDSSHGPGFFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLIKINSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSFLIVEEDAV ESSSEBSLGRSKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVHHSDTDALGYKEELGNIDPYPNIGG	166	1516		2010		000	
NKNMYTERLQNVMVLĒQCESDSSSLYREITY   SYLLAFNVWLLLAPVTLCYDWQVGSIPLVETI   WDMRNLATIFLAVVMALISIHCLAAFKRIE     HKEVIVGLIFLVFPFIPASNLFFEVGFVVAER     VLYMPSMGYCILFVHGISLLCTWLNRCGATT     LIVSTVLLILIFSWKTVKQNEIWLSRESILFRS     GVQTIPEINAKVHYNYANFILKDQGRNKEAIY     HYRTALINNIKAWDYLCWRFRKTLTDLP     1517	100	1310	A	2019	2	927	
SYLLAPNYWLLLAPYTLCYDWQVGSPLVETI   WDMRNLATIFLAVVMALLSIHCLAAFKRILE   HKEVLVGLIFLVFPFIPASNLFFRVGFVVAER   VLYMPSMGYCILPVHGLSKLCTWLNRCGATT   LIVSTVLLLLIFSWKTVKQNEIWLSRESLFFS   GVQTLPINAKVHYNYANFLKDQGRNKEAIY   HYRTALNNNKAWDYLCWRFRKTLTDLP     1517							
WDMRNIATIFLAVVMALLSLHCLAAFKRIE						,	
HKEVLVGLLFUNFTPIPASNLFFRVGFVVAER	l i						•
VLYMPSMGYCILFVHGLSKLCTWLNRCGATT   LIVSTVLLLLIFSWKTVK,ONEIWLSRESLFRS   GYQTLPHNAKVHYNYANFLKDQGRNKEAIY   HYRTALNNNKAWDYLCWRFRKTLTDLP     1517							
LIVSTVLLLLLFSWKTVKQNEIWLSRESLFRS GVQTLPHNAKVHYNYANFLKDQGRNKEAIY HYRTALNNKAWDYLCWRFRKTLTDLP  167 1517 A 2025 696 71 AAASAASSLTVTLGRLASACSHSILRPSGPGA ASLWSASRRINSQSTSYLPGYVPKTSLSSPPW PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY GRLFAVVHFASRQWKVTSEDLILIGNELDLA CGERIRLEKVLLVGADNFTLLGKPLLGKDLV RVBATVIEKTESWPRIIMFRRRKÑPKKRRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEETGQDQVLR PPRLRRAARVLMPLQAQVKAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEETGQDQVLR TVEISGEGGPLGHVVPFFSSLSGRILGLFIRGI EDNSRSKREGIFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEES VIGSLNIFONNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSHGPGPIFVKNILPKGGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKLINSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSEESLGESKEQLGDDVTKPSQKA  171 1521 A 2055 139 675 IPSRPWLQRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVHISDTDALGYKEPLGNIDFYPNGG							
GVQTLPHNAKVHYNYANFLKDQGRNKEAIY			l				
HYRTALNNNKAWDYLCWRFRKTLTDLP							
167 1517 A 2025 696 71 AAASAASSLTVTLGRLASACSHSILRPSGPGA ASLWSASRRPNSQSTSYLPGYYPKTSLSSPPW PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY GRLFAVVHFASRQWKVTSEDLILIGNELDLA CGERIRLEKVLLVGADNFTLLGKPLLGKDLV RVEATVIEKTESWPRIMRFRKRÄNFKKRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR VKQQAQTQTSEPQDLDQVPEEFQGQDQVLR TVEISGEGGPLGHTVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSQLKTA NLTGTDSPETDASASLQQNKSFRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH TVATTUSTENGT TVATTUSTENGT TVATTUSTENGT TVATTUSTENGT TVATTUSTENGT TVATTUSTENGT TVATTUSTENGT TVATTUSTENGT TVATTUSTENGT TVATTUSTENGT TVATTUST TV					' I		l
ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW PEVVLPDPVEETRIHHAEVVKKVNEMIVTGQY GRLFAVVHPASRQWKVTSEDLILIGNELDLA CGERRLEKVILVGADNFTILIGKPLLGKDLV RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAQVSAPKPLKGQAQAPQ RLQGAAVVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR 169 1519 A 2049 1 945 QNLEDREVINGVQTELTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH 170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG	167	1517	A	2025	696	71	
PEVVLPDPVEETRÌHAEVVKKVNEMIVTGQY GRLFAVVHFASRQWKVTSEDLILIGNELDLA CGERIRLEKVLLVQADNFTLLGKPLLGKPLLGKDLV RVBATVIEKTESWPRIIMRFRKRKNPKKKRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR  169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSQLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEPSLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGILPPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG							
GRLFAVVHFASRQWKVTSEDLILIGNELDLÄ CGERRIEKVILVGADNFTILGKPLLGKDLV RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVFEFQGQDQVLR KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQQ ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSEPSLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRTGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDPYPNGG							
CGERIRLEKVLLVGADNFTILGKPLLGKDLV RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV TTPQTVVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAQVSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR  169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQLIQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRTTGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG					1		
RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLIFGNINGONLGKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQILQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRTTGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG							
1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRAARVLMPLQAQVKAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEBSLGESKEQLGDDVTKPSSQKA  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEBSLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IFSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG				J	ļ		
168 1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR 169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH 170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSERSLGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG					l		:
RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR  169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEPSLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG	168	1518	A	2046	2	366	
PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR  169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTICETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG			[		j	1	
KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR  169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG					l	1	
169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG			1	I			
TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG	169	1519	A	2049	1	945	
EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH 170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG			1				
VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG		1	1		ł		
NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEPSLGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG			į		i	Í	QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS
SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG			j	ļ	J	J	VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA
SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG		·		ŀ	]		NLTGTDSPETDASASLQQNKSPRVPRLGGKPS
TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG		ļ		j		j	
SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG			ĺ	ſ	(	ſ	
ETASLVIARQEGHFLPRELVMFRSQSH     170			ŀ	ł			SGDRILEVNGRDVTGRTQEELVAMLRSTKQG
170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG							
VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG	170	1520	A	2050	363	1	
KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG		}	l	1	ļ		
ESSSÈESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG		1	ŀ		1		KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV
171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG			[		1		
DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG	171	1521	A	2055	139	675	
			·		1		
				Į	ŀ		LDQPGCPKTILGGFQYFKCDHQRSVYLYLSSL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  RESCTITAYPCDSYQDYRNGKCVSCGTSOKB
172	1522	A	2056	3	361 '	SCPLLGYYADNWKDHLRGKDPPMTKAFFDT AEESPFCMYHYFVDIITWNKNVR LIOHKSAVEYAOSHLSLVSMCKESHKCSEPK
	1323				30.	MEWKVKIRSDGTRYTTKRPVRDRILKERALKI KEERSGLTTDDDTMSEMKMGRYWSKEERKQ HLVRGKEQRRRREFMMRIRLKCLKES
173	1523	A	2060		387	GTRILSMQIPFVGFQPIRTSEHMAAAGVFALL QAYAFLQYLRDRLTKQEFQTLFFLGVSLAAG AVFLSVIYLTYTGYIAPWSGRFYSLWDTGYA KIHIPIIASVSEHQPTTWVSFFFDLHILGCTFPA G
174	1524	A	2071	74	443	LLMGPKAKKSGSKKKKVTKAERLKLLQEEEE RRLKEEEEARLKYEKEEMERLEIQRIEKEKW HRLEAKDLERRNEELEELYLLERCFPEAEKLK QETKLLSQWKHYIQCDGSPDPSVAQEMNT
175	1525	A	2083	139	486	AALTWSQPQEFWPMEMQPIVTDMVTVHWV AESSTVGWLCALFRVTHVGVGATGHGVVCG RRVLCGLPLPSPAPMPIMSLPEGESRKEREVQ RLQFPYLEPGHELPATTLLAFLAAV
176	1526	A	2092	3	587	EGSVNFKFGVLFAKDGQLTDDEMFSNEIGSEP FQKFLNLLGDTITLKGWTGYRGGLDTKNDTT GIHSVYTVYQGHEIMFHVSTMLPYSKENKQQ VERKRHIGNDIVTIVFQEGEESSPAFKPSMIRS HFTHIFALVRYNQQNDNYRLKIFSEESVPLFG PPLPTPPVFTDHQEFRDFLLVKLINGEKATLET PCI
177	1527	A	2103	44	427	GKGQVSLEGRPHRGPLCLGSWWPGSRVPGC CDGAWLAWACWVFGNDFPSPASAACSALLG CSVSTACLCVPLCSGSPLAPFRRTAALQEGLR RAVSVPLTLAETVASLWPALQELARCGNLAC RSDLO
178	1528	A	2104	2	409	ALQSTLGAVWLGLLLNSLWKVAESKDQVFQ PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD
179	1529	A	2111	1	312	PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG
180	1530	A	2116	3	366	TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV
181	1531	A	2117	2	386	YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVTTQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI
182	1532	A	2123	1	493	RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS
183	1533	A	2140	3	561	RQAWHEAPKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF VYGIDKFMEDLKDMLGFAPSRYYYYMWKYI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				soquance		SPLMLLSLLIASVVNMGLSPPGYNAWIEDKAS EEFLSYPTWGLAVCASLDVFAILPVPVAFIGR RFSLIDDGAGPFCSAAYTTTGCRTPYL
184	1534	A	2145	3	538	HELTVAAADRGQPPQSSVVPVTVIVLDVND NPPVFTRASYRVTVPEDTPVGAELLHVEASD ADPGPHGLVRFTVSSGDPSGLFELDESSGTLR LAHALDCETQARHQLVVQAADPAGAHFALA PVTIEVQDVNDHGPAFPLNLLSTSVAENQPPG TLVTTLHAIDGDAGAFGRLRYHL
185	1535	A	2151	2	671	LDKLLDRMENYNIFNEYILKQVAATYIKLGW PKNNFNGSLVQASYQHEELRREVIMLACSFG NKHCHQQASTLISDWISSNRNRIPLNVRDIVY CTGVSLLDEDVWEFIWMKFHSTTAVSEKKIL LEALTCSDDRNLLNRLLNLSLNSEVVLDQDAI DVIHVARNPHGRDLAWKFFRDKWKILNTRI RQKTLEFDFAEPLILAFPIILYTAIDNPPLVREH E
186	1536	A	2153	2	400	GPMCDKHSAFAEKFHAGFIDYIVHPLWETWA HLALPDAQDILYTLEDNRNWVDSMIPQSPSPP LDEQNRDWQGLLENLHVELTLDEEDSEGPEK EGEGQTYFTSSKTLCGIVPQNTDSLGETGIHIC AHDKSP
187	1537	A	2158	227	442	FNCFRVASDSFLENSSLLIMILPLRNATQEFIIR PGAVAYTCNPSTLGGWGGWITRSGVRDQPG OHGGTPS
188	1538	Α	2167	3	486	AHLGGAWLTQRSLGSWAAPGPARAAKEVVA CIPQNQKMNIWRMKTSKHLQLLSFVLGAVSP AVVVPYMMVLQENGYGVEEGIPTILIMAASS MDDILAITGFNTCLSIVFSSGCARSSGSRNSKS LRTPLGTICEGCDDSSIFSHLDHSSKWSSTYG HSGA
189	1539	A	2168	2	412	EFLSSNQITQLPNTTFRPMPNLRSVDLSYNKL QALAPDLFHGLRKLTTLHMRANAIQFVPVRIF QDCRSLKFLDIGYNQLKSLARNSFAGLFKLTE LHLEHNDLVKVNFAHFPRLISLHSLCLRRNKV AIVVSSLDW
190	1540	A	2179	64	399	MRLNONTLLLESFGXXRPYTSEHAPTYHOW MKADELLRWTTSEPLTLEHEYAMQRTWLED AYECTFIVLDAEKRHAQPGATEESCMVGDVN LFLTDLEDLTLGEIEVLIAEP
191	1541	A	2190	1	469	CLDRAAGIRHERNVIYINETHTRHRGWLARR LSYVLFIQERDVHKGMFATNVTENVLNSSRV QEAIAEVAAELNPDGSAQQQSKAVNKVKKK AKRILQEMVATVSPAMIRLTGWVLLKLFNSF FWNIQIHKGQLEMVKAATETNLPLLFLPVHR SH
192	1542	Α	2197	26	157	PSKXGGIRLLLTGTQLYGRFGSAIAPLGDLDR DGYNGEGREEPY
193	1543	A	2236	2	383	EYFPNSIWRSLFSTMDLGDIGFYTYRILQALS YTHSKGIMHRDVKPLNILCNSPRNKVILADW GLAEFYHPMRKYSVHVATRYYKSPEILLDYE YYDYSLDIWAVGVILLELLTLKLHVFEGGDN EQ
194	1544	A	2241	105	409	RKGVGKMPTSEGRPGQERSDWVTSYKVMGS NDSHTWVTVKNGSGDMIFEGNSEKEIPVLNE LPVPMGARYIRINPQSWFDNGSICMRMEILGC PLPDFNNY
195	1545	A	2245	1	672	MGVASDWTKRIEYQPGSGSMPLFPSIHLETCD GAVSSLQIVTELQTNYIGKGCDRETYSEKSLQ

Decide   Seq	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SST	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
Sequence			Į	,			F-recipialismine, G-Glycine, H-Histidine,
194			j			to lost amino	M=Methionine N=Asparasine P=Proline
maino acid residue   sequence	1 -	uonoc	1				O-Chitamine P-Arrinine S-Serine
Persidue of peptide   Pe	dano		<b>∤</b> .	7.7			TeThrennine VeVeline WeTentochen
peptide   pep		l	]				
	1.	ĺ		1		Sequence	
			ŀ				
			<del>                                     </del>			<del></del>	
MKHGRSFOVRAEKETILCYSDKTEMNRHHY   ALYYMPINCRUPILERDFOQADITRADEPHW   KLDQQALAKVDGQPGKSITRQLQEMPVTIQG   SILKPS	1			i	ĺ	1	MIFKEDGROGAKIPDGIVPKNI TDOFTTIMW
ALYYMNCRLYFILRRDFDQADITRPAREPHUN   KIDQQALAKVDGQPGKSITRQLQEMPVTIQG   SILKPS	1	1					MKHGPSPGVRAEKETII CYSDKTEMNRHHY
SILPQQALAKVDGQPGKSTTQLQEMPVTIQG   SILPS	1	ļ		ļ		ļ	
196	. '	Ì				}	
196				1			
197	196	1546	A	2256	1	396	
N.J.Y.IKTIPCTTRAPROGEV9GVDYNTSVEQF	1						
SALEESGALLESGTYDGNFYGTPKPPAEPSPF	1		İ				
197	1	ŀ		)			
197	1					,	
WLYKNPLGQSDNIPTIMFAISELTIPLAVICY	197	1547	Α	2259	43	594	
VKIIRTIDKTEIKRAFLAVSIALAINGVCTNTI KLIVGRPRDFFYRCFPOGVMNSEMECTGDP DLVSEGRKSFPSIHSSFAFSGLGFTTFYLAGKL HCFTESGRCKSWRLCAALLPL   198							
		l .			• .		
DLVSEGRKSFPSIISSFAFSGLGFTTFYLAGKL	1 1		[				KLIVGRPRPDFFYRCFPDGVMNSEMHCTGDP
1548	į l						DLVSEGRKSFPSIHSSFAFSGLGFTTFYLAGKL
PLGFASNNCFIMAAMSYDRYTAIHNPLQYHT	L						HCFTESGRGKSWRLCAAILPL
LIMTRICLQMMASSWMVGFLFSLCIIVTVPN   LSLCULNTIQHYFCDISPVVSLACNYTFYHEM   LSLCULNTIQHYFCDISPVVSLACNYTFYHEM   LSLCULNTIQHYFCDISPVVSLACNYTFYHEM   AIFVLSA	198	1548	A	2275	3	404	TCTTVVVIPRMLVDFLSESKTISLPECATQMFF
LIMTRICLQMMASSWMVGFLFSLCIIVTVPN   LSLCULNTIQHYFCDISPVVSLACNYTFYHEM   LSLCULNTIQHYFCDISPVVSLACNYTFYHEM   LSLCULNTIQHYFCDISPVVSLACNYTFYHEM   AIFVLSA							FLGFASNNCFIMAAMSYDRYTAIHNPLOYHT
1549							LMTRKICLQMMMASWMVGFLFSLCIIVTVFN
1549	[		·				LSLCDLNTIQHYFCDISPVVSLACNYTFYHEM
RHAAVLNNTVTAQIGIVAVVRGSLFFFPLELLI   KRLAFCHSNVLSHSYCVHQDVMKLAYADTL   PNVVYGLTALLVMGADRMFISLSYFLI							
	199	1549	Α	2315	1	375	
PNVVYGLTAILLVMGXDRMFISLSYFLII	1 1						
200							KRLAFCHSNVLSHSYCVHQDVMKLAYADTL
SDDPMLPSPDQLKKKAPFTNKKLKAHQTIPVD ILKQKAHQLASMQVQAYNGGNANPRPANNE EEEDEDEDYDYDYESLSDDNILEDRPENKSCH DQLQFEYKEEM  201 1551 A 2350 3 512 ISWEAQIAEIIQWVSDEKDARGYI.QALASKM TEELEALRSSSLGSRTLDPLWKVRRSQKLDM SARLELQSALEARRKQLVQEELRKVKDAN LTLESKLKDSEAKNRELLEEMEILKKKMEEK FRADTGKLML.CDSALFEYKYFSNECFYFLFD LIVITLEAPTEPQIQY  202 1552 A 2351 1 1003 PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDIILDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA CLAMLLHFLDTYQGLLQEEGAGHIIKDLYL LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPLFFHFRRIDSCLQTRVAFRGS DEEFCRYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVILQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EQVSPGODTEHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILINEEDMNVIVVD WSRGATTFTYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPGCLMSLSDLSLSPAPPSHLSPRCPSFQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T 205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGFWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC							
ILKQKAHQLASMQVQAYNGGNANPRRANNE EEEDEEDEYDYDYBSLSDDNILEDRPENKSCH DQLQFFYKEEM     201   1551   A   2350   3   512   ISWEAQIAEIIQWVSDEKDARGYLQALASKM TEELALRSSLGSRTLDPLWKVRRSQKLDM SARLELQSALEAGRAKQLVQEELRKVKDAN LTLESKLKDSEAKNRELLEEMEILKKKMEEK FRADTGKLMLCDSALFEYKYFSNECFYFLFD LIVTLEAPTEPQIQY     202   1552   A   2351   1   1003   PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA CLAMLLHFLDTYQGLLQEEGAGHIIKDLYL LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEFFCRYVMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPODTEHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE     203   1553   A   2361   2   403   NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP     204   1554   A   2390   280   476   SPSLLFQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T     205   1555   A   2400   543   745   AAVALRDISWQQPYPMDFYAGSSLGFWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	200	1550	A	2334	2	409	
201 1551 A 2350 3 512 ISWAQIAERIQWVSDEKDARGYLQALASKM TELEALRSSLGSRTLDPLWKVRRSQKLDM SARLELQSALEAERRAKQLVQEELRKVKDAN LTLESKLKDSEAKNRELLEMEILKKKEEK FRADTGKLMLCDSALFEYKYFSNECFYFLFD LIVTLEAPTEFQIQY  202 1552 A 2351 1 1003 PSSYSSDELSPGPLTSPPWAPLGAPERPEHIL NRVLERALGGATRDSAASDILDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA CLAMLLHFLDTYQGLLQEEGAGHIKDLYL LIMKDESLYQGLREDTLRHQLVETVELKIPE ENQPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEFAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHILFACTRDSYEALVPLPE EIQVSPGDTEHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYPPPGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFTYNRAVKNTRKVAVSLSVHIKNL LKHQASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSILPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRCAREMDATFMPPAPSCPSERV T AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC							
DQLQFEYKEEM	] ]						
201 1551 A 2350 3 512 ISWEAQIAEIIQWVSDEKDARGYLQALASKM TEELEALRSSSLGSRTLDPLWKVRRSQKLDM SARLELQSALEAEIRAKQLVQEELRKVKDAN LTLESKLKDSEAKNRELLEEMEILKKKMEEK FRADTGKLMLCDSALFEYKYFSNECFYFLFD LIVTLEAPTEFQIQY  202 1552 A 2351 1 1003 PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQVFVRAGGMEGPEGLGRKQA CLAMILHFLDTYQGLLQEEGAGHIKDLYL LIMKDESLYQGLREDTIRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQVSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVIFNTQKKTVWLIH GYRPVGSIPLWLQNFFVRILLNEEDMNVIVVD WSRGATTFTYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	l i		1				
TEELEALRSSSLGSRTLDPLWKVRRSQKLDM SARLELQSALEAEIRAKQLVQEELRKVKDAN LTLESKLKDSEAKNRELLEEMEILKKKMEEK FRADTGKLMLCDSALFEYKYFSNECFYFLFD LIVTLEAPTEFQIQY  202 1552 A 2351 1 1003 PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA CLAMLLHFLDTYQGLLQEEGACHHIXDLYL LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRIDSCLGTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPODTEHRVEEPDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFTYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T 205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGFWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	201	1551	<b>—</b> —	2250		510	
SARLELQSALEAEIRAKQLVQEELRKVKDAN LTLESKLKDSEAKNRELLEEMEILKKKMEEK FRADTGKLMLCDSALFEYKYFSNECFYFLFD LIVTLEAPTEFOIQY  202 1552 A 2351 1 1003 PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA CLAMLLHFLDTYQGLLQEEGAGHIKDLYL LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFRTQKKTVWLIH GYRPVGSPLWLQNFVRILLNEEDMNVIVVD WSRGATTFTYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T AVAIRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	201	1331	^	2330	3	312	
LTLESKLKDSEAKNRELLEEMEILKKKMEEK FRADTGKLMLCDSALFEYKYFSNECFYFLFD LIVTLEAPTEFQIQY  202 1552 A 2351 1 1003 PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA CLAMLLHFLDTYYGLLQEEEGAGHIKDLYL LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSGGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPODTEIHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFTYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC		i					
FRADTGKLMLCDSALFEYKYFSNECFYFLFD LIVTLEAPTEPQIQY  202 1552 A 2351 1 1003 PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA CLAMLLHFLDTYQGLLQEEEGAGHIKDLYL LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPJFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFTYNRAVKNTRKVAVSLSVHIKNL LKHQASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	<b>!</b>		1				SAKLELŲSALEAEIKAKŲLVŲEELKKYKDAN
LIVTLEAPTEFQIQY	) }						ED ADTOVI M. COCAL ESTANCEUR CONTROLLER
202 1552 A 2351 1 1003 PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA CLAMLHFLDTYQGLLQEEGAGHIKDLYL LIMKDESLYQGLREDTIRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFTYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGFWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	]						
NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA CLAMLLHFLDTYQGLLQEEGAGHIKDLYL LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	202	1552	$\overline{\mathbf{A}}$	2351	1	1003	
LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA CLAMILHFLDTYQGLLQEEGAGHIIKDLYL LIMKDESLYQGLREDTI.RLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFTYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC		1002	^	1,01	*.	1003	
CLAMILHFLDTYQGLLQEEGAGHIKDLYIL LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEFRDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC				ļ			
LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEFEDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFTYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	1	ĺ	{	ì		'	CLAMITHE DANGE LOSSEC & CRIBEDTA
ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEHRVEPRDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNIL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC							LIMKDEST AOOI BEDJA DI HOT ABLADA BADDA
DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC					.		
TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEPBDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP LGRITGLDP AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC		i			ł		
PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEPBDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHQASLDNFHFIGGSLGAHISGFVQKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	1			ļ	l	·	
EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE  203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVQKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC				1			[ ·
203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	1 1	1	1		į		
203 1553 A 2361 2 403 NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC			]				
GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	203	1553	A	2361	2	403	
WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	-				-	.~-	
LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP  204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC			İ				
204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T 205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC					l		
204 1554 A 2390 280 476 SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC			ŀ			ł	
AGSRLGAMRRCAREMDATPMPPAPSCPSERV T  205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	204	1554	$\overline{\mathbf{A}}$	2390	280	476	
205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	1						
205 1555 A 2400 543 745 AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC			i				
HGQDRRPHAPGRPARGKVQEGSARPPSAVAC	205	1555	$\overline{\mathbf{A}}$	2400	543	745	
			ŀ				
	[		Ì				EDCSCR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	]	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
J		1	]	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
		L	<u> </u>	sequence		nucleotide insertion
206	1556	Α	2406	122	485	DLSPDSREDHPQGHRRLLPKRPVRGSLMPGH
}	,	ļ	)		•	THHPCPVSSTTNDTPDQIWVSVGSLRMGTGG
	ļ	i	İ	•		MGANASTSPRCWDLSSGNKKWIIQVPILASIV
		L	<u> </u>			ESRGGLLATGVGGMCACVPRNQPLTGT
207	1557	Α	2409	289	418	LWILYRHKQQVQHNHSNRLSCRPSQEDRAT
	1550		L			HTIMVLDKENTLS
208	1558	A	2413	64	492	VQGTGXXFIAFTEAMTHFPASPVWAGMFFL
ļ			1		;	MLINLGLGSMIGTMAGITTPIIDTFKVPKEMFT
İ		1		1		GGCCVFAFLVGLLFVQRSGNYFVTMFDDYSA
]	]				j	TLPLTLIVILENIAVAWIYGTKKFMQELTEML
	1550					GFRPYRFYFYMWKFVSP
209	1559	Α	2417	3	877	EKERLLDEWFTLDEVPKGKLHLRLEWLTLMP
j	l .	j			j	NASNLDKVLTDIKADKDQANDGLSSALLILY
			1			LDSARNLPIRYKTNEPVWEENFTFFIHNPKRQ
Ì		ļ		·		DLEVEVRDEQHQCPLGNLKVPLSQLLTSEDM
			1			TVSQRFQLGNSGPNSTIKMKIALRVLHLEKRE
		1				RPPDHQHSAQVKRPSVSKEGRKTSIKSHMSG
j		ļ	j	]		SPGPGGSNTAPSTPVIGGSDKPGMEEKAQPPE
			]			AGPQGLHDLGRSSSSLLASPGHISVKEPTPSIA
,		1	l			SDISLPIATQELRQRLRQLENGTTLGQSPLGQI OLTIP
210	1560	A	2422	35	456	
210	1300	^	2422	33	430	REFAASDLEPFTPTDQPISPEAITQPSCIKRQRA AGNPGSLAATIDHKPCSAPLEPKIOASRNORW
ł I	•	}	l '			GAVRAAESLTDIAEPASPQVHETPIDASQTQK
i I	ļ	l	]			VEPASKSRFTPELQAKVSHSRERALSTMDATP
		1	}			HHAOPORGEG
211	1561	A	2431	1	764	RRYSOKLIOHTACOLLRTYPAATRIDSSNPNP
	****	1 **		• 00	,04	LMFWLHGIQLVALNYQTDDLPLHLNAAMFE
ł	ł					ANGGCGYVLKPPVLWDKNCPMYQKFSPLER
		l				DLDSMDPAVYSLTIVSGONVCPSNSMGSPCIE
	]	İ				VDVLGMPLDSCHFRTKPIHRNTLNPMWNEOF
	]		,			LFHVHFEDLVFLRFAVVENNSSAVTAORIIPL
		1				KALKRGYRHLQLRNLHNEVLEISSLFINSRRM
	1	İ				EENSSGNTMSASSMFNTEERKCLQTHRVTVH
		1				GVPG
212	1562	Α	2436	1	411	GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI
		I	}	ŀ		YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC
	·	1				TOLDHNLDKGYLTVLGEQATPTNRLGALPKG
		l	Ì			RANRTRDLELTYLAERIVRLTWIPGDANNRPI
		L		_		TDYDCQIEEHQ
213	1563	Α	2445	1	1294	MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT
						WLMPELHPKEQILELLVLEQFLSILPEELQIWV
						QQHNPESGEESVTLLEDLEREFDDPGQQVPAS
	1	1			ĺ	PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ
	}	l				LKSWKPCLSPKSDCENSETATKEGISEEKSQG
						LPQEPSFRGISEHESNLVWKQGSATGEKLRSP
						SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT
		l		ļ		TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT
		l	1			QHQRIHTGEKPYKCNQCGKAFSLRSYLIIHQR
		ľ				IHSGEKAYECSECGKAFNQSSALIRHRKIHTG
						EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY
		1	1			ECNECGKTFSQSSKLIRHQRIHTGERPYECNE
		1				CGKAFRQSSELITHQRIHSGEKPYECSECGKA
						FSLSSNLIRHQRIHSG
214	1564	Α	2461	1	615	GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL
		l			•	STRLVSVQEDAGKSPARNRSASITNLSLDRSG
						SPMVPSYETSVSPQANRTYVRTETTEDERKIL
1						LDSVQLKDLWKKICHHSSGMEFQDHRYWLR
L			L ˈ			THPNCIVGKELVNWLIRNGHIATRAQAIAIGQ
	·	·		L		

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ł		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l	ł	peptide		/=possible nucleotide deletion, \=possible
				sequence	L	nucleotide insertion
						AMVDGRWLDCVSHHDQLFRDEYALYRPLQV
215	1565	A	2464	3	2932	LFSVYCQLECSKLIL
213	1303	^	2404	3	2932	GPGVRSSQDGMADVFVHLRTAWPRCSFISGQ HGPGRHGRRVCSSQDSMADVFVHLRTAWPT
						CSLISGQHGPGESVSYEDDDIPAPASLLHVNA
						AAPALTNPTAPVLCTAPNNTAQKEKVPSGMR
1					ļ	QRPAGVRISSRTPDLTCAVSTHSTVPGVRISSC
1					}	TPDLTCAVSIHSTVPSVCISSCTPDLTCAVSTH
1					. '	STVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
						TPDLTCAVSIHATVPGVRISSCTPDLTCAVSIH
				·		ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
						TPDLTCAVSIHSTVPGVRISSCTPDLTCAVSIH ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
						TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH
						ATVPGVRISSCTPDLTCAVSIHATVPGVRISSC
				·		TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH
1 1					·	ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
		ľ				TPDLTCAVSIHATVPGVRISSCTPDLTCAVSTH
}						STVPGVRISSRTPDLTCAVSIHATVPGVHISSC
					i	TPDLTCAVSTHSTVPGVRISSRTPDLTCAVSIH
						STVPGVCISSRTPDLTCAVSIHSTVPSVHISSCT
1						PDLTCAVSIHSTVPGVRISSRTPDLTCAVSTHS TVPGVHISSCTTDLTCAVSIHATVPGVHISSCT
						PDLTCAVSTHITVPGVRISSRTPDLTCAVSIHS
						TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRT
						PDLTCAVSTHLTVPGVRISSRTPDLTCAVSIHA
1						TVPGVHISSCTPDLTCAVSIHATVPGVRISSRT
1 1						PDLTCAVSIHATVPGVHISSCTPDLTCAVSTHS
				į		TVPGVRISSRTPDLTCAVSIHSTVPGVHISSCT
						PDLTCAVSTHSTVPGVHISSCTPDLTCAVSTH
1 1					·	STVPGVHISSRTPDLTCAVSIHATVPSVHISSC TPDLTCAVSIHSTVPGLLTSVSQTSTG
216	1566	A	2477	1	414	FRTKSYRKGSYRCIVSEWIAEQGNWQEIQEK
				· j	''-'	AVEVATVVIQPTVLRAAVPKNVSVAEGKELD
		- 1				LTCNITTDRADDVRPEVTWSFSRMPDSTLPGS
		1				RVLARLDRDFLVHSSPHVALSHVDARSYHLL
						VRDVSKENSGYYY
217	1567	Α	2480	2	460	CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY
					1	MQVLVCQHECVRELATRPGRLSPIENFLPLHY
[ [		ľ	i	1	ľ	DYLQFAYYRVGEYVKALECAKAYLLCHPDD
] ]	ļ	١	}			EDVLDNVDYYESLLDDSIDPASIEAREDLTMF VKRHKLESELIKSAAEGLGXSYTEPNYW
218	1568	$\overline{\mathbf{A}}$	2483	140	383	AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT
			2.03	- "	J.J.J	SANLLQLVRSSGDIQEGDLVEVVLSASATFED
	×					LQIRPHALTVHSYRAP
219	1569	A	2489	3	428	SSRLVLLAGAAALASGSQGDREPVYRDCVLQ
	ĺ	ļ				CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD
'			l	j		DCKYECMWVTVGLYLQEGHKVPQFHGKWP
[ ]	1	- 1	ľ	ľ	1	FSRFLFFQEPASAVASFLNGLASLVMLCRYRT
1220	1,600	l	0466			FVPASSPMYHTCVAFAWVS
220	1570	A	2498	1	1297	MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA
( l		Į	ļ	1	. 1	APKIPRLVQATPAFMAVTLVFSLVTLFVVDH
			ł	•	.	HHFGREAEMRELIQTFKGHMENSSAWVVEIQ
	}	l	j		ļ	MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL
ļ		ļ		l	. [	KEDLEKADALTFOTLNFLKSSLENTSIELHVL
j		]	}	ļ	- 1	SRGLENANSEIQMLNASLETANTQAQLANSS
1 1	i	[	ĺ	1	ţ	LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS
L				İ	]	LEGANAEIQGLKENLQNTNALNSQTQAFIKSS
				·		

SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	,,,,,,,	in in	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	F=Phenylalamine, G=Glycme, H=Hishdine,   I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Scrine.
		<b>}</b> .	717	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		i		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ.	İ	peptide	Soqueto	/=possible nucleotide deletion. \=possible
	]	j		sequence		nucleotide insertion
			-			FDNTSAEIQFLRGHLERAGDEIHVLKRDLKM
1			i			VTAQTQKANGRLDQTDTQIQVPKSEMENVN
	]	}	]	Ì		TLNAQIQVLNGHMKNASREIOTLKOGMKNA
1	[	1		ſ		SALTSOTOMLDSNLOKASAEIORLRGDLENT
				Ì		KALTMEIQQEQSRLKTLHVVITSQEQLQRTQ
221	1571	A	2501	3	500	RVRLNNDGLSPLMMAAKTGKIGIFOHIIRREV
l	[	[	1	[ -	555	TDEDTRHLSRKFKDWAYGPVYSSLYDLSSLD
1	ł	1	j	1		TCGEEASVLEILVYNSKIENRHEMLAVEPINE
		1				LLRDKWRKFGAVSFYINVVSYLCAMVIFTLT
	1	1			Ì	AYYQPLEGTPPYPYRTTVDYLRLAGEVIILFT
L		1		ļ		GVLFFFTN
222	1572	A	2508	3	395	DAHCQRKLAMQEFMEINERLTELHTQKQKL
						ARHVRDKEEEVDLVMQKVESLRQELRRTER
						AKKELEVHTEALAAEASKDRKLREQSEHYSK
1			<b>[</b>	ļ		QLENELEGLKQKQISYSPGVCSIEHQQEITKL
L			L			KTDLEKKS
223	1573	À	2544	2	412	NDPALISNESAAVVHTIVNETLESMTSLEVTK
l '			1			MVDERTDYLTKSLKEKTPPFSHCDQAVLQCS
1						EASSNKDMFADRLSKSIIKHSIDKSKSVIPNID
						KNAVYKESLPVSGEESQLTPEKSPKFPDSQNQ
204	1671		0000	403		LTHCSLSAA
224	1574	A	2552	401	1	GASLCFISTAFTVLTFLIDSCRFSYPERPIIFLSM
						CYNIYSIAYIVRLTVGRERISCDFEEAAEPVLI
						QEGLKNTGCAIIFLLMYFFGMASSIWWVILTL
						TWFLAAGLKWGHEAIEMHSSYFHIAAWAIPA
225	1575	A	2563	724	1	MSARKERREKGEEEGEGEKDGDEDEKEEEKE
		**				GLGEEEKEAGKKKKKQEEKEKEKGAVYSR
1 1						VARICKNDMGGSORVLEKHWTSFLKARLNC
						SVPGDSFFYFDVLQSITDIIQINGIPTVVGVFTT
						QLNSIPGSAVCAFSMDDIEKVFKGRFKEOKTP
1					·	DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK
]						TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT
						KTRVRYRLTAISVDHSAGPYH
226	1576	A	2571	449	3	EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP
						TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV
					-	AGAACDRGMSLEACEAVTRKANRRTYTMG
			Ì			VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE
L	1.676		0505			RGVIREKMMPADAIVDHIMDRIFS
227	1577	A	2575	3	1197	VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS
						EDHINRKYAFKAAHPNMRTYYFCTDTGKEM
				_		ELWMKAMLDAALVQTEPVKRVDKITSENAP
ļ ·						TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE
1	{		'		•	KKALEAEKYGFQKDGQDRPLTKINSVKLNSL
]						PSEYESGSACPAQTVHYRPINLSSSENKIVNVS
1						LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ
[	[				1	LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP
}						STHDKTLGPGAEEKRRSMRDDTMWQLYEW
.				l		QQRQFYNKQSTLPRHSTLSSPKTMVNISDOT
<b> </b>	1			l	1	MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI
				İ	·	QRGDVTIDRRHRAHHPKVK
228	1578	A	2583	3	330	LPFLGLGSVLPQGMVMASPEMNPTICSVFEA
		**	2000	-		HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE
j j	<b> </b>	j				KIHAKVRGTWPFISPEVRKEGGLPOTGRELLD
					· [	PTMGIKPHLWWVAA
229	1579	A	2589	1	448	DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR
] ]	· · ·			-		ECVAPNICKCKPGYIGSNCQTALCDPDCKNH
ļ						GKCIKPNICQCLPGHGGATCDEEHCNPPCQH

SEQ ID	SEQ ID	Met ·	SEQ	Predicted	Predicted end	I Andrew Market Control
NO: of	NO: of	hod	ID NO:			Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in No.	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	]	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		]	914	ng to first	acid residue	Q=Ghutamine, R=Arginine, S=Serine,
	1	ł	, , ,	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
j	ļ	}	ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ı		peptide	bodeomen	possible nucleotide deletion, possible
}	1	i	1	sequence	7	nucleotide insertion
						GGTCLAGNLCTCPYGFVGPRCETMVCNRHC
	]	ļ		<u> </u>		ENGGQCLTPDICQCKPGWYGPTCSTA
230	1580	A	2593	2	138	AVIFSVVFAYVADITQEHERSMAYGLVCMFI
1	1	İ		{ <sup>-</sup>	1	LYLLYLLRNAFFLR
231	1581	A	2595	185	2	SGPYTDFTPWPTEEQKLLEQALKTYPVNPPER
1	l	i	i		-	WEKIAEAVPGRTKKACIKRYKVADLRISK
232	1582	A	2596	1	391	STVTGQPRRLLDTAGHQQPFLELKIRANEPGA
1	ľ	í		ì		GRARRTPTCEPATPLCCRRDHYVNFQELGW
ļ	,					RDWILLPEGYQLNYCSGQCPTHLAGSPGIAAS
f	ĺ	ĺ	İ I	1		FHSAVFSLLKANNPWPGRTSWCVPTARRPLS
L		1	,			LLYL
233	1583	Α	2601	184	403	LLFSDEIIMAAPLRIADVTSGLIGGEDGRVYV
						YNGKETTLGDMTGKCKSWITPCPEEKVNVLQ
L						NSIPYWERIT
234	1584	A	2614	178	335	PLTLCLPENNKPPQADAVPDKELTLPVDSTTL
						DGSKSSDDQKIISYLWEKTQ
235	1585	A	2616	2	896	DVLEVYGTGVASTRHEMGTLDKHKELEDLV
	1					AKFLNVEAAMVFGMGFATNSMNIPALVGKG
ł	•					CLILRDEVNHTSLVLGARLLGATIGIFKHNYA
[						QSLEKLLRDAVIYGQPRTRRAWKKILILVEGV
1						YSMEGSIVHLPQHALKKKYKAYLYIDEAHSI
				•	•	GAVGPTGRGVTEFFGLDPHEVDVLMGTFTKS
1						FGASGGYIAGRKARILSPPACLVPNTGSHSLH
					-	RLTRDLQMNEAMVALVTDRLQGWNSGEGN
						WDRADKFGDLVDYLRVHSHSAVYASSMSPPI
236	1506		0.001			AEQIIRSLKLIMGLDGTTQ
236	1586	Α .	2621	1	392	NTSSFPAQPSSPARPSLPHLSQHPSNPLLPLAS
ļ						ADHPQCGRFLPLHEPEPLCPSPSLSYPTLVSS
						WSSPFSSHHGCPPGLYPFPTSPKTIQPPGLAQL
						KMLCIPPGRQQLRGAQSMPGHGALSPLLLPP
237	1587	A	2628	398	1	DLVCKISGFGRGPRDRSEAVYTTMSGRSPAL
]	.50,	^	2026	376	1	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF
	ĺ		1			GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN
1		1	j			LMHRLMLDCWQKDPGERPRFSQIHSILSKMV
		ĺ	ĺ			ODPEPPNV QADFOERFRESQIESILSKMV
238	1588	A	2631	i	1104	WSPCSLTCGVGLQTRDVFCSHLLSREMNETV
_				-	1101	ILADELCRQPKPSTVQACNRFNCPPAWYPAQ
1	1	ł	ł			WQPCSRTCGGGVQKREVLCKORMADGSFLE
	i	İ				LPETFCSASKPACQQACKKDDCPSEWLLSDW
]	- 1	ı	ŀ	1		TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS
	1	İ	1	1		TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI
	ſ	1	i		ļ.	AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA
		i	l	ł		VVLRCPARRVRKPLITWEKDGQHLISSTHVT
		J	J	j	ļ	VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF
	ļ	j	ł	į		VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP
-	ĺ	j	J	l	j	KEALOTHKHONGIFSNGSKAEKRGLAANPGS
		l			ľ	RYDDLVSRLLEQGAPCSSSKKKN
239	1589	A	2636	1	678	MKPDNILLDEHGHVHITDFNIAAMLPRETQIT
				-	***	TMAGTKPYMAPEMFSSRKGAGYSFAVDWW
· [	ľ	ı	l	Ì		SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET
,		. [	1		1	TVVTYPSAWSQEMVSLLKKLLEPNPDQRPSQ
	i	- 1	ł		- 1	LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK
	ĺ	. [				GRLNCDPTPELEEMILESKPLHKKKKRLAKK
' l	ł	1	l	l	ł	EKDMRKCDSSQTCLLQEHLDSVQKEFIIINRE
	ŀ	I			•	KVNRDCI
240	1590	A	2639	389	3	ELLDPTTPMRTKCIELLYAALTSSSTDQPKAD
ľ					-	LWQNFAREIEEHVFTLYSKNIKKYKTCIRSKV
- 1		1	1	}	j	ANLKNPRNSHLQQNLLSGTTSPREFAEMIVM
	L				<del></del>	MA LIMINATIN 101 I DOTTILI AND IN THE THE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghtamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Ghtamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  EMANKELKQLRASYTESCIQEHYLPQVIDGTL Y
241	1591	A	2640	392	3	IRLTILRCVFMRLATICVLVFTLGSKITSCDDD TCDLCGYNQKLYPCWETQVGQEMYKLMIFD FIIILAVILFVDFPRKLLVTYCSSCKLIQCWGQ QEFAIPDNVLGIVYGQTICWIGAFFSPLLPAM Y
242	1592	A	2642	405	1	YFKNTTLLLVGVICVAAAVEKWNLHKRIALR MVLMAGAKPGMLLLCFMCCTTLLSMWLSNT STTAMVMPIVEAVLQELVSAEDEQLVAGNSN TEEAEPISLDVKNSQPSVELIFVNEDILDFLMK SPLMISQACI
243	1593	A	2646	412	2	CLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAF LLNGSIDGIRHMFTFKLEIMLEPKVWREAATQ VFFALGLGFGGVIAFSSYNKRDNNCHFDAVL VSFINFFTSVLATLVVFAVLGFKANVINEKCIT ONSETV
244	1594	A	2650		1271	MTTTLIGLLKTARLLRLVRVARKLDRYSEYG AAVLMLLMCIFALIAHWLACIWYAIGNVERP YLTDKIGWLDSLGQQIGKRYNDSDSSSGPSIK DKYVTALYFTFSSLTSVGFGNVSPNTNSEKIF SICVMLIGSLMYASIFGNVSAIIQRLYSGTARY HMQMLRVKEFIRFHQIPNPLRQRLEEYFQHA WTYTNGIDMNMVTNGTCSSCTSDDGHFILVS NHHQGGLIYSWNDAASMQRPFNHIKSSLLGS TSDSNLNKYSTINKIPQLTLNFSEVKTEKKNSS PPSSDKTIIAPKVKDRTHNVTEKVTQVLSLGA DVLPEYKLQAPRINKFTILHYSPFKAVWDWLI- LLLVIYTAIFTPYSAAFLLNDREEQKRRECGY SCSPLNVVDLIVDIMFIIDILINFRTTYVNQNEE VVSDPASV
245	1595	A	2656	385	2	NLTWWPLFRDVSFYIVDLIMLIIFFLDNVIMW WESLILLTAYFCYVVFMKFNVQVEKWVKQ MINRNKVVKVTAPEAQAKPSAARDKDEPTLP AKPRLQRGGSSASLHNSLMRNSIFQNKIHTLD PHV
246	1596	A	2660	200	506	VLVLQMNYYQMLIIYYVLFFKVNEFLAFEGPI LLDMRIKHLIKTNQLSQATALAKLCSDHPEIG IKGSFKQTYLVCLCTSSPNGKLIEEVSMFSFIS NYFLS
247	1597	A	2678	3	267	DAWVKNDIIFNQTERKQKISENLKHLASVRV VQKNLVFVVGLSQRLADPEVSPLVFFVILIFF VSLSYLEIIFDPAQLCDSSEHIIS
248	1598	A	2687	1	404	DFTTLAAMMRTLFSLFGDVRSDVHRFSVTLF GAAIKSVKNPDKKSIENQVLDSLVPLLLYSQD ENDAVAEESRQVLTICAQFLKWKLPREVYSK DPWHIKPTEAGTICRFFEKKCKGKINILEQTL MYSKNPKL
249	1599	Ā	2692	1	440	FRRRRRRERDCAAQGARRHCRHLAECKLV SFPIGIYKVLRNVSGQIHLITLANNELKSLTSK FMTTFSQLRELHLEGNFLHRLPSEVSALQHLK AIDLSRNQFQDFPEQLTALPALETINLEENEIV DVPVEKLAAMPALRSINL
250	1600	A	2693	459	21	LLPGSLGVPILHSQPWDPSPQCPHRAPSTPRRL PPLGALSQALTFLSRAAKNHSQDPGKGTKPFP AAPAAPPPRSSLPAPLPMGLKDKGPQPAPPTIF NSPWHPATLPGALGPQLSQAAPSPIPPPCLMG ISSCPDLKLTKSSTP
251	1601	Α	2694	2	404	FVFDLKLRVPGFAALLIHGASSVPGPETVRLR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	
nucl-	peptide	liou	in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid,
cotide	seq-		USSN			F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence		09/496	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
uence	ucite			correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
nence	ì		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		-	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			ſ	peptide		/=possible nucleotide deletion, \=possible
	ļ		<u> </u>	sequence		nucleotide insertion
1						QKRKKKAPDHSSGRKEELVTTHTVDKLETKK
						PVGRVLCGLSGELLHSLLLPRRKTEKRALGSH
[	[	1		[	ľ	RKAGFPEHPVAPEPLSNSCQISKEGREQVLSEI
L				1	İ	GAGDCL
252	1602	A	2697	421	1	PQKSHSGAYQCFATRKAQTAQDFAIIALEDG
	ļ	l				TPRIVSSFSEKVVNPGEQFSLMCAAKGAPPPT
	ì	ì				VTWALDDEPIVRDGSHRTNQYTMSDGTTISH
		1	1	j		MNVTGPQIRDGGVYRCTARNLVGSAEYQARI
}		j .	J i	i	1	NVRGPPSIRAMRNIT
253	1603	A	2698	65	401	ACCOWRTLIPAKSTTVSCTISTPHHPFRGSYS
	1000	^	2070	03	401	
		1			1	FDDHITDSEALSRSSHVFTSHPRMLKRQPAIEL
		1		•	l	PLGGEYSSDVPRPLSTQLSSSLLGYFSTLMTG
054	1.00				l	AAFTNNIASSTIIL
254	1604	A	2699	438	301	GQIHSQDDPPFIDQLGFGVAPGFQTFVACQEQ
				-		RVRGPWEAGPGVGY
255	1605	Α	2700	1	842	LQNREDSSEGIRKKLVEAEELEEKHREAQVS
[ [		l				AQHLEVHLKQKEQHYEEKIKVLDNQIKKDLA
		1				DKETLENMMQRHEEEAHEKGKILSEQKAMIN
1 1		•				AMDSKIRSLEQRIVELSEANKLAANSSLFTQR
1						NMKAQEEMISELRQQKFYLETQAGKLEAQN
] ]						RKLEEQLEKISHODHSDKNRLLELETRLREVS
1						LEHEEQKLELKRQLTELQLSLQERESQLTALQ
1 1			1			
1			i i		ĺ	AARAALESQLRQAKTELEETTAEAEEEIQALT
256	1606	A	2701	2	405	VGLGSNIFRLLKASARMSVELALSILAHP
200	1000	А	2/01	2	403	FVGGPGADPPVAVMWDPRAARMDLTAYAB
1			]			LLKESGNQVLKNGNFSLAIRKYDEAIQILLQL
1						YQWGVPPRDLAVLLCNKSNAFFSLGKWNEA
						FVAAKECLQWDPTYVKGYYRAGYSLLRLHQ
<u> </u>						PYEAARMFFEGLR
257	1607	A	2702	2	399	FVESASSRPPGCFSGDGRFWLVSEGSRRGWD
l i						FNPSFSFLDPRYSVGGDENIGTVTTLANILREF
[ [						NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE
l i				·		DLPVQARRLVDLMKNDTRIHFQEDWKIITLFI
1 1						GGNDL
258	1608	A	2709	1	1097	SVGARQGEARDRIRRFFPKGDLEVLQAOVERI
j ,				-	1007	MTRKELLTVYSSEDGSEEFETTVLKALVKACG
	•					
]		l				SSEASAYLDELRLAVAWNRVDIAQSELFRGDI
		ļ	ļ <b>[</b>			QWRSFHLEASLMDALLNDRPEFVRLLISHGLS
} }		- 1				LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH
		l			l	SAGTKAPALKGGAAELRPPDVGHVLRMLLG
j l						KMCAPRYPSGGAWDPHPGQGFGESMYLLSD
[	j	(	[	ĺ	ľ	KATSPLSLDAGLGQAPWSDLLLWALLLNRA
]		l	1	· 1		QMAMYFWEMGSNAVSSALGACLLLRVMAR
ĺĺ		- 1	•	- 1		LEPDAEEAARRKDLAFKFEGMGVDLFGECYR
	1			j		SSEVRAARLLLRRCPLWGDATCLQLAMQAD
L		}			j	ARAFFAQDGVQSLPTQKWWGDMARR
259	1609	A	2721	1	403	VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR
	Į	J		· ]	-	GGFSQREMVTGERSPSPEEEEEEEEEGFGERA
[ <b> </b>		i		}		SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA
•		l		ŀ		PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE
	ĺ	ĺ	ļ	ĺ	Ĭ	
260	1610	<del></del>	2720	<del></del>	477	GGLRVRLP
200	1010	A }	2728	1	477	LLGGDLRYHLQQNVHFTEGTVKLYICELALA
	1		j	I	ı	LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN
	į		ł	j	j	IATVVKGAERASSMAGTKPYMAPEVFQVYM
	ļ	l	İ	ŀ	. ]	DRGPGYSYPVDWWSLGITAYELLRGWRPYEI
l			}	į		HSVTPIDEILNMFKVERVHYSSTWCKGMVAL
	[	[	1	ľ	ľ	LRK
261	1611	A	2730	3	547	LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR
				ļ		EIYLQGCFKPLVSISPNDSLFEAVYTLIKNRIH
						HINNIH I VEGETABLISHED TO THE TOTAL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	}		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	1	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \-possible
			<u> </u>	sequence		nucleotide insertion
	i	ĺ	1		}	RLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLP
			İ			RPSFLYRTIQDLGIGTFRDLAVVLETAPILTAL
		i	Į			DIFVDRRVSALAVVNECGTHPQDERLGLGW
			<u>.                                    </u>			GLGEPGSEERLFPAAITSR
262	1612	Α	2733	3	431	GPEFPGSAKLVFLDLSYNNLTQLGAGAFRSA
		[		ĺ	(	GRLVKLSLANNNLVGVHEDAFETLESLOVLE
	}	İ	'			LNDNNLRSLSVAALAALPALRSLRLDGNPWL
	i	ł	ł '		i	CDCDFAHLFSWIQENASKLPKGLDEIQCSLPM
	·					ESRRISLRACRRPASRV
263	1613	A	2736	2	343	PARISGVDPPVRKATKGGENCSFEDNKNWOF
						LWGLNGNFNFFKEPWGGRNNHAKGFRTTW
						ARSSSQNNRTFQNNRNFLRLQRDSQKKGQFA
						RLISPLVNLPQSPGGLEFQYQAT
264	1614	A	2738	2	245	RAMLKCLREGQPPPSYNWTRLDGPLPSGVRV
				_		DGDTLGFPPLTTEHSGIYVRHDTNEFSSRDSH
'			1			DTVDVLDPPEDSGKOVDL
265	1615	A	2752	2	388	AAGDAPLRSLEQANRTRFPFFSDVKGDHRLV
	1010		2.02	*	300	LAAVETTVLVLIFAVSLLGNVCALVLVARRR
						RRGATACLVLNLFCADLLFISAIPLVLAVRWT
			1 !			EAWLLGPVACHLLFYVMTLSGSVTILTLAAV
				i i		SLER
266	1616	A	2755	192	1	AFREVGGYWGLLCEHLYAIPSKTSEGNWTAK
200	1010	Λ.	2155	192	1	
						LQGYLPLQDAFHIFQDPLTGDLPWPELILGLP
267	1617	A	2760	434	714	
207	1017	^	2700	434	/14	ASRLEKQNSTPESDYDNTPNDMEPDGMGYM
						HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF
268	1618	$\overline{\mathbf{A}}$	2762	1	405	LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF
200	1010	Α .	2/02	i	400	IACIFCGQDEWSPERSTRCFRRSRFLAWGEP
						AVLLLLLLSLALGLVLAALGLFVHHRDSPL
- 1		- 1	i 1			VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP
1						ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE
200	1710		0000			LPLSWAE
269	1619	A	2772	3	243	TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHL
		- 1				LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN
						LSKNLSFSEFCFDVSY
270	1620	Α	2789	1	486	ELQSQQACTHTKETEQLRSQLQTLKQQHQQA
- 1		- 1				VEQIAKAEETHSSLSQELQARLQTVTREKEEL
l		l	. 1			LQLSIERGKVLQNKQAEICQLEEKLEIANEDR
		J				KHALERFEQEAVAVDSNLRVRELQRKVDGIQ
İ						KAYDELRLQSEAFKKHSLDLLSKERELNGKL
						RHLSP
271	1621	A	2795	1	568	KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE
[	Į	ſ	ĺ	ſ	. [	FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE
	i	l	l	.		RKHEHLKNKSAPKVVKQKVIDAHLDSOTON
ł	ŀ	- 1	}		1	FOOTOIOTAESKAEHKKLPOPYNSLOEEKCLE
ł		l	l	ļ		VKGIOEKOVFSNTKDSKOEITONKSFFSSVKE
		J	J	1	<b></b> ]	SQRDDGKGALNIVEFLRKREELHQILSTVKQP
272	1622	$\overline{\mathbf{A}}$	2797	8	523	KCMQGKYAGAMESEPCVCTEADFDCDYGYE
		l		- J		RHSNGQCLPAFWFNPSSLSKDCSLGQSYLNST
i	}	ľ	ł		l	GYRKVVSNNCTDGVREQYTAKPOKCPGKAP
		ļ				RGLRIVTADGKLTAEQGHNVTLMVOLEEGD
l	ļ		ļ		ł	
		ĺ	j	Į.		VQRTLIQVDFGDGIAVSYVNLSSMEDGIXHV
273	1622	$\overline{}$	2001	<del></del>	205	YQNXGIXRXTVQVDNSLGS
213	1623	Α	2801	72	395	HPSRSNVGPRQLTVWNTSNLSHDNRRKYIFS
		!		l	-	DEEGQNQLGIRIHQDIPLPPRRRELPALRTTNG
	}	1		1	1	KADSLNVSRNSVMQELSELEKQIQVIRQELQL
- 1	1		1			AVSRKTELEEYH
	1 40 4					
274	1624 .	Α	2805	168	320	ILWLYFETGTWVYPVFAKLSLLGLAALFSLRE IFIARNGVVGETLTHCKRV

NO: of   NO: of   NO: of   NO: of   No: of   N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Decision   Decision						nucleotide	
Sequence	nuci-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
uence	eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
amino acid residue of peptide sequence   Ti-Threonie, V-Valina, W-Tryptoplan, Sequence   Peptide sequence   Peptide sequence   Peptide sequence   Peptide	seq-	uence	}	09/496	correspondi		
residue of peptide sequence	uence			914			
Popsible   Popsible   Popsible   Popsible mucleotide deletion, Popsible   pulcebide insertion   Popsible   pulcebide insertion   Popsible   pulcebide insertion   Popsible   P	ł	l	l	ł			
		i	Į.	1		sequence	
275	ł						, ·
MGKILFQ			L .	<u> </u>		L	
276	275	1625	A	2812	208	321	
KYQQQTVVADILAGDETIPGSSLLPGHVQA	L			<u> </u>			
277   1627   A   2817   3   410   VLQERLDNFQRKCQLASSTEGKVDKLLMF   LFISYLHTPKIKQHEVLQAMGSILGITGEEN   PLFQEHGTATRWMTGWLGGGSKYPKTY   GLNQPALNGSFSELFKKTESLSSTLFT   GLNQPALNGSFSELFKKTESLSSTLFT   LPFINSFGKIK   GLSOPSCSCPHSPLFTIJSRAQLETALKWRN   VKLRLLHLEELQMEHDIRHYDLESVFMTV   VKLRLLHLEELQMEHDIRHYDLESVFMTV   VKLRLLHLEELQMEHDIRHYDLESVFMTV   VKLRLLHLEELQMEHDIRHYDLESVFMTV   PVDQNFRLV   PLPANLFAHSNPLQPLPSLPHFFLPATHKFP   TPFITFSSVPPPLPSLSSILHHSPLHSELMPHLC   CRLPSRFSVSRELPPGASSVPLAPTPLPD   VPSQRIPTTXPFPAS   VFSQRIPTTXPFPAS   VSSQRIPTTXPFPAS   VSSQRIPTTXPFPAS   VSSQRIPTXPFPAS   VSSQRIPTXPFPAS   VSSQRIPTXPFPAS   VSSQRIPTXPFPAS   VSSQRIPTXPFPAS   VSSQRIPTXPFPAS   VSSQRIPTXPFPAS   VSSQRIPTXPFAS   VSSQRIPTXPSYSSPGLSPFTSSCSFSNPD   VSSQRIPTXPSSSCSFSNPD   VSSQRIPTXPSSSCSFSNPD   VSSQRIPTXPSSSCSFSNPD   VSSQRIPTXPSSSCSFSNPD   VSSQRIPTXPSSSCSFSNPD   VSSQRIPTXPSSSCSFSNPD   VSSQRIPTXPSSSCSFSNPD   VSSQRIPTXPSSSCSFSNPD   VSSQRIPTXPSSCSSCSFSNPD   VSSQRIPTXPSSCSFSNPD   VSSQRIPTXPSSCSSSCSFSNPD   VSSQRIPTXPSSCSSCSFSNPD   VSSQRIPTXPSSCSSCSSSCSSCSSCSSCSSCSSCSSCSSCSSCS	276	1626	A	2813	41	266	
1627							
LFISYLHTPKIRKQHEVLQAMGSLGITGEER    PLPQEHGTATRWMTGWLEGGSKSVPKTP    GLNQQPALNGSFSELFVKFLKTESLSSTLPT    CPHNSFGKIK			ļ.,	<u> </u>	<u> </u>		
PLFQEEHGTATR WMTGWLEGGKSSYPKTP   GLNQQPALNGSFSELFVKFLKTESLSSTLFT   GLNQQPALNGSFSELFVKFLKTESLSSTLFT   GLNQQPALNGSFSELFVKFLKTESLSSTLFT   LPPINSPOKIK   CLSGFSCSCPHSPLFTISRAQLETALKWRN7   VKLRILLHLEELQMEHDIRHYDLESVPMTV   VKULLLHLEELQMEHDIRHYDLESVPMTV   VVQNPRLV	277	1627	A	2817	3	410	
GLNQQPALNGSFSELFVKFLKTESLSSTLPT	١.		1				
LPPHNSPGKIK		ĺ	ĺ		ĺ ·		
278							
VKLRILLH.EELQMEHDIRHYDLESVPMTV	279	1620		2021	220	457	
PVDQNPRLV	2/0	1028	^	2021	236	[ 437	
1629							•
TPPTFSSVPPPLPSI.SBILHISPLHSELNPHI.C   CRLPSRPSVSRELPPQSGPASSVPLAPTPLPD   VPSQRIPTXPPPAS	279	1620	Δ	2822	342		
CRLPSRPSVSRELPPQSGPASSVPLAPTPLPD	213	1029	1 ^	2022	342	1	
VPSQRHPTKFPPAS							
280							VPSORHPTXPPPAS
CGQYWPLEKDSRRFGFLTVTNLTGAVGEP	280	1630	A	2825	307	77	
281   1631   A   2827   81   381   KMGTAVWYKEKEKDKASQEGGDVLGA QDCTPSLKSLVATGNLLDLEETAKAPLSTV: NTTNMDEVPRPQALSGSSVVWVSGCVASR VILSLTSG   LISTSG   KLPXDKYELEPSPLTQYILERKSPHTCWQVI TSSGKYNELGYPFGYLKASTTLTCVNLFVM YNYPVLLPLLDDLFKVHKLKPNLKWRQAF YLKTLPPYYL   VSPALSTPTIFSYSPSPGLSPFTSSSCFSFNPJ MKHYLHSQACSVFNYHLSPRTFPRYPGLMM PLQCQMHPESTQFSIKLQPPPVGRKNRERV SSEESSP   LIST   L					1 22.		
281							
NTTNMDEVPRPQALSGSSVVWVSGCVASR	281	1631	A	2827	81	381	KMGTAVWVPKEKEKRDKASQEGGDVLGAR
VILSLTSG		l	ļ	l	þ		QDCTPSLKSLVATGNLLDLEETAKAPLSTVSA
1632				]			NTTNMDEVPRPQALSGSSVVWVSGCVASRS
TSSGKYNELGYPFGYLKASTTLTCVNLFVM							VILSLTSG
YNYPVLLPLDDLFKVHKLKPNLKWRQAFI   YLKTLPPYYL	282	1632	A	2830	471	160	KLPXDKYELEPSPLTQYILERKSPHTCWQVFV
YLKTLPPYYL							TSSGKYNELGYPFGYLKASTTLTCVNLFVMP
1633		ļ	ļ				YNYPVLLPLLDDLFKVHKLKPNLKWRQAFDS
MKHYLHSQACSVFNYHLSPRTFPRYPGLMT	-	4.55	<u> </u>				<u> </u>
PLQCQMHPEESTQFSIKLQPPPVGRKNRERY   SSEESAP	283	1633	A	2835	462	148	
SSEESAP		[	ĺ	[			
284 1634 A 2836 2 384 KTLPRTLLDILADGTILKVGVGCSEDASKLL DYGLVVRGCLDLRYLAMRQRNNLLCNGLS KSLAETVLNFPLDKSLLLRCSNWDAEILTE QVIYAARDAQISVALFLHILGYPFSRNSPGE KR  285 1635 A 2843 20 271 PIRPYYSYSGLDRDCSWLPLAKAWLPDVMI VCDRVSEDGINRQQAQEWCIKHGFELVELS EELPEEDGKCLCVRRYGTYI  286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFINSELHRA NLHVGNLRLLSGPLDQVRAALPTPALSPKD AVLQNLKRILAK VQEMRDQR VSLEQQLREI QKDDITGSLVTIDHSQMKLFEEQLKKYDC KVYLEQNLAAQDRVLCALT  288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELD LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYITAKYVDHRFSRKTCSTSS KLNELLEAIKSRDLLALIQVYAEGVELMEPI			1				
DYGLVVRGCLDLRYLAMRQRNNLLCNGLS KSLAETVLNFPLDKSLLLRCSNWDAEILTE QVIYAARDAQISVALFLHLLGYPFSRNSPGE KR  285 1635 A 2843 20 271 PIRPYYSYSGLDRDCSWLPLAKAWLPDVMI VCDRVSEDGINRQQAQEWCIKHGFELVELS EELPEEDGKCLCVRRKYGTYI  286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFINSELHRA NLHVGNLRLLSGPLDQVRAALPTPALSPKD AVLQNLKRILAKVQEMRDQRVSLEQQLREI QKDDITGSLVTIDHSQMKKLFEEQLKKYDK KVYLEQNLAAQDRVLCALT  288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELD LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYTTAKYVDHRFSRKTCSTSS KLNELLEAIKSRDLLALIQVYAEGVELMEPI	204	1624	-	2026	<del></del>	204	
KSLAETVLNFPLDKSLLLRCSNWDAEILTE QVIYAARDAQISVALFLHLLGYPFSRNSPGE KR  285 1635 A 2843 20 271 PIRPYYSYSGLDRDCSWLPLAKAWLPDVMI VCDRVSEDGINRQQAQEWCIKHGFELVELS EELPEEDGKCLCVRRKYGTYI  286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFINSELHRA NLHVGNLRLLSGPLDQVRAALPTPALSPKD AVLQNLKRILAKVQEMRDQRVSLEQQLREI QKDDITGSLVTTDHSQMKKLFEEQLKKYDK KVYLEQNLAAQDRVLCALT  288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELD LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYTTAKYVDHRFSRKTCSTSS KLNELLEAIKSRDLLALIQVYAEGVELMEPI	204	1034	^	2030	<b>4</b>	304	
QVIYAARDAQISVALFLHLLGYPFSRNSPGE KR  285 1635 A 2843 20 271 PIRPYYSYSGLDRDCSWLPLAKAWLPDVMI VCDRVSEDGINRQQAQEWCIKHGFELVELS EELPEEDGKCLCVRKYGTYI  286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFINSELHRA NLHVGNLRLLSGPLDQVRAALPIPALSPKD AVLQNLKRILAKVQEMRDQRVSLEQQLREI QKDDITGSLVTTDHSQMKKLFEEQLKKYDK KVYLEQNLAAQDRVLCALT  288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELD LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYTTAKYVDHRFSRKTCSTSS KLNELLEAIKSRDLLALIQVYAEGVELMEPI			i				
KR			ł				
285 1635 A 2843 20 271 PIRPYYSYSGLDRDCSWLPLAKAWLPDVMI VCDRVSEDGINRQQAQEWCIKHGFELVELS EELPEEDGKCLCVRKYGTYI  286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK  287 1637 A 2851 2 427 FVAEVREWAKYMEVHEKASFTNSELHRA NLHVGNLRLLSGPLDQVRAALPTPALSPKD AVLQNLKRILAKVQEMRDQRVSLEQQLREI QKDDITGSLVTTDHSQMKKLFEEQLKKYDC KVYLEQNLAAQDRVLCALT  288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELD LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYTTAKYVDHRFSRKTCSTSSKLNELLEAIKSRDLLALIQVYAEGVELMEPI		Ĭ	i				
VCDRVSEDGINRQQAQEWCIKHGFELVELS EELPEEDGKCLCVRKYGTYI  286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVREWAKYMEVHEKASFTNSELHRA NLHVGNLRLLSGPLDQVRAALPTPALSPKD AVLQNLKRILAKVQEMRDQRVSLEQQLREI QKDDITGSLVTTDHSQMKKLFEEQLKKYDC KVYLEQNLAAQDRVLCALT  288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELD LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYTTAKYVDHRFSRKTCSTSS KLNELLEAIKSRDLLALIQVYAEGVELMEPI	285	1635	A	2843	20	271	
286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFINSELHRA NLHVGNLRLLSGPLDQVRAALPTPALSPKD AVLQNLKRILAKVQEMRDQRVSLEQQLREI QKDDITGSLVTTDHSQMKKLFEEQLKKYDC KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELD LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYTTAKYVDHRFSRKTCSTSS KLNELLEAIKSRDLLALIQVYAEGVELMEPI	<b>-</b>		l	1	l		
286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFINSELHRA NLHVGNLRLLSGPLDQVRAALPTPALSPKD AVLQNLKRILAK VQEMRDQR VSLEQQLREI QKDDITGSLVTIDHSQMKKLFEEQLKKYDC KVYLEQNLAAQDR VLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELD LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYITAK YVDHRFSRKTCSTSS KLNELLEAIKSRDLLALIQVYAEGVELMEPI		}	l	ĺ			EELPEEDGKCLCVRRKYGTYI
287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFINSELHRA NLHVGNLRLLSGPLDQVRAALPTPALSPKD AVLQNLKRILAKVQEMRDQRVSLEQQLREI QKDDITGSLVTTDHSQMKKLFEEQLKKYDC KVYLEQNLAAQDRVLCALT  288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELD LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYITAKYVDHRFSRKTCSTSS KLNELLEAIKSRDLLALIQVYAEGVELMEPI	286	1636	A	2845	197	278	<del>1</del>
NLHVGNLRLLSGPLDQVRAALPTPALSPKD AVLQNLKRILAKVQEMRDQRVSLEQQLREI QKDDITGSLVTTDHSQMKKLFEEQLKKYDC KVYLEQNLAAQDRVLCALT  288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELD LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYITAKYVDHRFSRKTCSTSS KLNELLEAIKSRDLLALIQVYAEGVELMEPI							FVAEVRREWAKYMEVHEKASFINSELHRAM
AVLQNLKRILAKVQEMRDQRVSLEQQLREI QKDDITGSLVTTDHSQMKKLFEEQLKKYDC KVYLEQNLAAQDRVLCALT  288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELD LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYITAKYVDHRFSRKTCSTSS KLNELLEAIKSRDLLALIQVYAEGVELMEPI			1				NLHVGNLRLLSGPLDQVRAALPTPALSPKDK
QKDDITGSLVTTDHSQMKKLFEEQLKKYDO KVYLEQNLAAQDRVLCALT  288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELD LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSKLNELLEAIKSRDLLALIQVYAEGVELMEPI		1	l	l			AVLQNLKRILAKVQEMRDQRVSLEQQLRELI
288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELD: LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYITAKYVDHRFSRKTCSTSS KLNELLEAIKSRDLLALIQVYAEGVELMEPI			l ·	1			QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL
LGTSELLPAKNVGNNSFNDIMEANLPSPSPK TPSSDMTVRKEYITAKYVDHRFSRKTCSTSS KLNELLEAIKSRDLLALIQVYAEGVELMEPI			l .	1			
TPSSDMTVRKEYITAKYVDHRFSRKTCSTSS KLNELLEAIKSRDLLALIQVYAEGVELMEPI	288	1638	A	2859	2	469	FVNLGILTCIECSGIHREMGAHISRIQSLELDK
KLNELLEAIKSRDLLALIQVYAEGVELMEPI			Ì				LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP
		l	l	!			TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA
EPGQELAETALHLAVRTADQTSLHLVE		ł	ł	1	}	}	KLNELLEAIKSRDLLALIQVYAEGVELMEPLL
		<u></u>	<u> </u>	L	<u> </u>		
	289	1639	A	2861	2	454	FVASGGPATARMSDSQFFCVAEERSGHCAVV
DGNFLYVWGGYVSIEDNEVYLPNDEIWTYI			l		ł		DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI
DSGLWRMHLMEGELPASMSGSCGACINGK		Į.	l	1			DSGLWRMHLMEGELPASMSGSCGACINGKL
		ľ	i	ľ	·		YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK
ITDFEGQPPTPRDKLSCWVYKDRLIYFG		<u></u>	L	L			
	290	1640	A	2868	1	378 .	FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI
		ł	1				SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF
PDCASCLQAQDPLCGWCVLQGRCTRKGQC	ł	l	L	<u> </u>		L	PDCASCLQAQDPLCGWCVLQGRCTRKGQCG

Mode   Mode	SEQ ID	SEQ ID	Met	SEQ	Predicted	I be at a deal of	
			1			Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence			200				DEASPARTIC ACID, EEGRITAMIC ACID,
Sequence		1	ł	1			
URRICE	1		ł				
minio acid residue of peptide residue of peptide residue of peptide sequence   for peptide   for peptide sequence   for peptide sequence   for peptide sequence   for peptide sequence   for peptide sequence   for peptide   for peptide sequence   for p			l				O=Chromina P=Assinina C=Corina
Peptide   Sequence			Į.				TeThreonine VerValine WeTryntonhon
Peptide	ļ		Į	1			
		1		i			
RAGQINQWI-WSYEEDSHCLHIQSLIPGHHPR			l				
291						<del>                                     </del>	
1641				1			
	291	1641	Α	2870	1	385	
	1	1	1	1	1		
1642   A   2877   3   188				Ì	Ì		
1642   A   2877   3   188	1						
PPPPAYYSSRYVAVICHGML VSCWCHI    293	L			1			
1643	292	1642	A	2877	3	188	RPTRPPPATTQSPESTMDTSLKKEKSAILDLYI
GAGTHPDAADPSGERTCOSEGSRSVLDLVNYF   LSPEKLTAEDRYYCESCASLQDAGKVYLLSQ   CPCYLLILLRFSFDLRTMRRKILDDVSIPLL   LRLPLAGGROQAYDL   CRCFFFGTTLLVYLLSFIGMVIFTFTLDLRYI   INVENTIGATION   CRCFFFGTTLLVYLLSFIGMVIFTFTLDLRYI   INVENTIGATION   CRCFFFGTTLLVYLLSFIGMVIFTFTLDLRYI   INVENTIGATION   CRCFFFGTTLNYLLSFIGMVIFTFTLDLRYI   INVENTIGATION   CRCFFFGTTLNYLLSFIGMVIFTFTLDLRYI   INVENTIGATION   CRCFFFTLNTNY   FINCELPGARKTIDAPPSLQPFLQDSKYNTALS   LASSQHGILNNISLLFSICKTCRTMDHHCPRA   NNCVGEQNHRFFCALHCKSKHFCTEFTLNTNY   FINCELPGARKTIDAPPSLQPFLQDSKYNTALS   LSESISQ   SQYSHSLDYHLLQVTKNPFTLGDSSNPGQTE   RLQEFSQKMDQVRGHWPVST   SPETILALDTFTLLGQDNILVLLLAFPFMAGG   SLXTSTMGRTRLRNKNPACEMAVVLLANLA   QGDSLAARAIAVQKGSIGHLIGFLEDSLAAT   QQQSQASLLIMHNPPFETSVDMMRRACRA   LLALAKVDDNHSSF   SULINASAQVNL   STATLLALAKVDDNHSSF   SULINASAQVNL							
CAGTHPDAAPSGERTCCSEGRSKUDLVNYE	293	1643	A	2878	1	427	REKEEEVEEEDKVVKETEKEAEQEKEEDSL
GPCYLLITLIRFSFDLRTMRRRKILDDVSIPLL   LRIPLAGGRQAYDL							
LRIPLAGGROQAYDL	'	1	ſ	i .		ĺ	LSPEKLTAENRYYCESCASLQDAEKVVELSQ
1644		1					
1645			1				LRLPLAGGRGQAYDL
295	294	1644	A	2879	109	245	
NNCVGEQNIHRFFCALHCKSKHFCIEFTI.NTNF   FNCTLPGAEKSTIIDAPPSI.QPFLQDSKYNTALS   LSESISQ							
FINCELPGARKSTIDAPPSLQPFLQDSKYNTALS   LSESISQ	295	1645	J A	2880	3	320	
LSESISQ			ľ				
296		i·					
RLQEFSQKMDQVRGHWPVST						<u> </u>	
1647	296	1646	A	2892	209	363	SQYSHSLDYHLLQVTKNPFTLGDSSNPGQTE
	202	1.647		2000		45.1	RLQEFSQKMDQVRGHWPVST
QGDSLAARAIAVQKGSIGHLLGFLEDSLAAT   QiQQQASLLHMinppfeptsvDmmrracra   Lialakvddninesef	291	1047	A	2893	8	424	SPXTLXLDTFILLGIQDNILVLILATPPFMAGG
LLALAKVDDNHSEF		Ï					
1648							QIQQSQASLLHMHNPPFEPISVDMMRRACRA
1649	298	1648	Δ	2804	210	145	
1649		10.0	1.	2074	310	443	
GYFQAYNVLILTMQASLPKVLRFCACAGMIY   LGYTFCGWIVLGPYHDKFENLNTVAECLFSL   VNGDDMPATFAQIQQKSILVWLFSRLYLYSFI   SLFYMILSLFIALITDSYDTIKKFQQNGFPETD   LQEF	299	1649	A	2898	1	492	KIKAKNI TNYDI CSIFI GTSTI I VWVGVIDVI
LGYTFCGWIVLGPYHDKFENLNTVAECLFSL   VNGDDMFATFAQIQQKSILVWLFSRLYLYSFI   SLFTYMILSLFIALITDSYDTIKKFQQNGFPETD   LQEF					- 1		
VNGDDMFATFAQIQQKSILVWLFSRLYLYSFI   SLFTYMILSFIALITDSYDTIKKFQQNGFPETD   LQEF	•						LGYTFCGWIVI GPYHDKFENI NTVAECLESI
SLFTYMILSLFIALITDSYDTIKKFQQNGFPETD   LQEF							
LQEF							SLFIYMILSLFIALITDSYDTIKKFOONGFPETD
TVTVRFVNKADFPKVRAKEQTFMFPENQPVS   SLVTITIGSSLRGEPMSYYIASGNLGNTFQIDQ   LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP   FSSYEKLDITVLDVNDNAPIF					ļ		
TVTVRFVNKADFPKVRAKEQTFMFPENQPVS   SLVTITIGSSLRGEPMSYYIASGNLGNITFQIDQ   LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP   FSSYEKLDITVLDVNDNAPIF	300	1650	A	2901	1	445	PVWWNSLNGASEVTFSVHVKDGGSFPKTDST
SLVTTTTGSSLRGEPMSYYIASGNLGNTFQIDQ		j		]	ļ		TVTVRFVNKADFPKVRAKEQTFMFPENOPVS
LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP   FSSYEKLDITVLDVNDNAPIF     301							SLVTTTTGSSLRGEPMSYYIASGNLGNTFOIDO
SSYEKLDITVLDVNDNAPIF   301   1651   A   2902   162   433   THFICLPLGYCFPLLDKDLQLPSGFNCNFDFLE   EPCGWMYDHAKWLRTTWASSSSPNDRTFPG   KPAVSEDMKELRPACSTYFNPRFPYKL   302   1652   A   2909   2   412   GPQMLCKKIYFTWVTRSQCQFEWLADIMQEV   EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI   CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN   SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ   LVNRQDRAHFM   SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ   LVNRQDRAHFM   VPPTSILEHLQRRKIMKRPSSCS   304   1654   A   2926   179   354   PGVPSQALRKAESLKKCLSVMEAKVKAQTAP   NKDVQREIADLGEVGAASLPPSSGPGA   305   1655   A   2938   135   438   GMGYLHAKGILHKDLKSKNVFYDNGKVVIT   DFGLFSISGVLQAGRREDKLRIQNGWLCHLA   PEIIRQLSPDTEEDKLPFSKHSDVFALGTTWYE   LHAREWP   306   1656   A   2944   2   329   VRWNSCVNCSCAFGNGASLSTSLGESSGCLW   EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA							1 ma 0 1 1 m 1 m 1 m 1 m 1 m 1 m 1 m 1 m 1 m
BPCGWMYDHAKWLRTTWASSSSPNDRTFPG   KPAVSEDMKELRPACSTYFNPRFPYKL     302							
BPCGWMYDHAKWLRTTWASSSSPNDRTFPG   KPAVSEDMKELRPACSTYFNPRFPYKL     302	301	1651	Α	2902	162	433	THFICLPLGYCFPLLDKDLQLPSGFNCNFDFLE
302 1652 A 2909 2 412 GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV EENDHQDLVSVHIYVTQLAEKFDLRTIMLYI CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM  303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA NKDVQREIADLGEVGAASLPPSSGPGA GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA					.		
BENDHQDLVSVHIYVTQLAEKFDLRTTMLYI CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM  303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS  304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA							
BENDHQDLVSVHIYVTQLAEKFDLRTIMLYI CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM  303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	302	1652	Α	2909	2	412	<b>GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV</b>
CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM  303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA			1				
LVNRQDRAHFM		ļ	- 1	j		ļ	CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN
LVNRQDRAHFM			l	1		ļ	SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ
VPPTSILEHLQRRKIMKRPSSCS  304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA  305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTTWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA							LVNRQDRAHFM
304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTTWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	303	1653	Α	2914	291	453	KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE
NKDVQREIADLGEVGAASLPPSSGPGA  305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTTWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA		l				<u></u>	
305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTTWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	304	1654	A	2926	179	354	PGVPSQALRKAESLKKCLSVMEAKVKAQTAP
DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTTWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA			!				
PEIIRQLSPDTEEDKLPFSKHSDVFALGTTWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	305	1655	Α	2938	135	438	GMGYLHAKGILHKDLKSKNVFYDNGKVVIT
306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	-	ļ	ļ		-	1	
306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA		- 1	i	ļ	j		PETIRQLSPDTEEDKLPFSKHSDVFALGTTWYE
EIGKWLSCSLLSFPSPLAVLITTFCIVTVLGREA			لــــبـــا				LHAREWP
	306	1656	A	2944	2	329	
LTKGALWAVFLLAGSALLCAEVTGVIWRQPE	1	1				}	
					1		LTKGALWAVFLLAGSALLCAEVTGVIWRQPE

NO: of   No: of   N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino said recuence (AmAlania C. C. d.)
mucleotide seq							Amino acid sequence (A=Alanine C=Cysteine,
Sequence	1						
Sequence	eotide		1				
1662   A   2967   3   465   A   2967   3   465   A   2967   3   465   A   2967   3   466   A   2967   4   2967   4   2967   4   2   33   1   1   1   1   1   1   1   1	seq-						
amino scid residue   sequence	uence	l		914			O=Glutamine, R=Arginine, S=Serine.
Persidue of peptide   Sequence   Y=Tyrosine, X=Unknown, **Sipp codon, / Pepsible nucleotide delicion, \				1		of peptide	
peptide	ł		1			sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
SITELLSFKVSSSA	ļ	1			peptide		/=possible nucleotide deletion, \=possible
1657					sequence		
POLSTTEGSHAFI PCX_ARGSPEPNITWDKDGQ			<u> </u>				
PySGAEGKFTIQPSGELLVKNLEGQDAGTY	307	1657	A	2950	2	411	
				ļ	· ·	İ	
		1	1	İ	ĺ		PVSGAEGKFTIQPSGELLVKNLEGQDAGTYT
1658		j	١.	<b>[</b>		İ	
DSSLPEEEDEDKEAINGGRAENRERHSESS   DWMKTYPSTINQTNSSMDFRNYMMRDETLEP   LPKNNFEMA TJTDTGMYFFDINTKTTTWLDP   LPKNNFEMA TJTDTGMYFFDINTKTTTWLDP   RLCKKAKAFEDC   DPLTLTLTEPTGLLYVGAREALFAFSMEALE   LQGAVRGGAVGGSRACQARFRGAVLG     310   1660   A   2959   1   419   QDMMERAIDITYOHDVERSVYSVAMTIQ   HVAREHFRILEFWMKTMGLINAVIPVAWPT   TOFVQLSISVTALTAILKYGQVLMHSHVVIW   LPLAVYAVATIMECP   TOFVQLSISVTALTAILKYGQVLMHSHVVIW   LPLAVYAVATIMECP   LAVYAVATIMECP   LPLAVYAVATIMECP   LPLAVYAVATIMECP   LPLAVYAVATIMECP   RWFQMFGLGAFNGVFGGFRR RVAFFLLFFSSROYLGVMKAQKFGAGGGKK   POKFFGLRGTLKFQKSGGHEIBGAWBGAFGFNGGLEPQ   KI   RWFQMFGLGAFNGVFGGGFRR RVAFFLLFFSSROYLGVMKAQKFGAGGGKK   POKFFGLRGTLKFQKSGGHEIBGAWBGAFGFNGGLEPQ   KI   RWFQMFGLGAFGFNGGLEPGAWBGAFGFNGGLEPGAWBGAFGFNGGLEPGAWBGAFGFNGGLEPGAWBGAFGFNGGLEPGAWBGAFGFNGGLEPGAWBGAFGFNGGLEPGAWBGAFGFNGGLEPGAWBGAFGFNGGLEPGAWBGAFGFNGGLEPGAWBGAFGFNGGLEPGAWBGAFGFNGGLEPGAWBGAFGFNGGLEPGAWBGAFGFNGGLEFGAWBGAFGFNGGLEFGAWBGAFGFNGGLEFGAWBGAFGFNGGLEFGAWBGAFGFNGGLEFGAWBGAFGFNGGLEFGAWBGAFGFNGGLEFGAWBGAFGFNGGLEFGAWBGAFGFNGGLEFGAWBGAFGFNGGLEFGAWBGAFGFNGGLEFGAWBGAFGFNGGAFGAFGFNGGAFGAFGFNGGAFGGAF	308	1658	A	2951	1	407	
DWMKTYPSYNQTNSSMDFRNYMMRDETLEP   LPKNWEMAYJDTGMYFDRNITKTTTWLDP   RLCKKAKAPEDC   GPFLTLTLTEYTGLIYVGAREALFAFSMEALE   LQCAVRGGAVGGSRACQRARRGAVLG   LQCAVRGGAVGGSRACQRARRGAVLG   LQCAVRGGAVGGSRACQRARRGAVLG   LQCAVRGGAVGGSRACQRARRGAVLG   LQCAVRGGAVGGSRACQRARRGAVLG   LQCAVRGGAVGGSRACQRARRGAVLG   LQCAVRGGAVGGSRACQRARRGAVLG   LQCAVRGGAVGGSRACQRARRGAVLG   LQCAVRGGAVGGSRACQRARRGAVLG   LQCAVRGGAVGGSRACQRARRGAVLG   LTAVYAVATIMFCF   LTAVYAV	""	1030	<b>'</b>	2731	[ •	707	Deel Definition of the Principle of the
LIPKNWEMAYIDIGMIYFIDENTIKTITULIP		ì			1		
RICKKAKAPEDC	l		ļ	ļ		į	
1699   A   2954   2   179	Ī		1	Į			
1660	309	1659	A	2954	2	179	
1661	310	1660	Α	2959	1	419	QDMMERAIIDTFVGHDVVEPGSYVQMFPYPC
HIVAEKEHRIKEVMKTMGLNNAVHWVAWPIT TGFVQLSISVTALTALKGQVLMHSHVVIIW							YTRDDFLFVIEHMMPLCMVISWVYSVAMTIQ
1661							HIVAEKEHRLKEVMKTMGLNNAVHWVAWFI
1661							TGFVQLSISVTALTAILKYGQVLMHSHVVIIW
1662	211	1661		20.62		422	
PQKPGLRGTILKPQKSGHOHENOPPGPCNA RVAPMLPRLIPPOPSDKEGGWGLKSQPPS	211	1001	A	2963	3	465	
RVAPMLLPRLPTTGVPSDKEGGWGLKSQPF8							
AVQNGKLPGHQPPNGYGPGAEPGFNGGLEPQ   XI			. '				
1662	. •						
1662							
HTGGCSPKTKPHIKEECIYPTPCYKPKEKLPY EAKLPWFKQAQELEEGAAVSEEPSFIPEAWS ACTVTCGVOTQVRIVRCQVILSFSQSVADLPI  313 1663 A 2969 2 430 VVADNCRQGYLDALRFLERRGLTKEPVLWT LVSKEPPAPADGNWDAGCDQRRKGGLSLNW KVPHVQVKDVPNFEQLSPELEAALKKACTRD PSRWARFWISGPGQVLTYLLLPCTLPFEYIYF RSRRLVVWLPDVPADLWWMQ  314 1664 A 2971 422 33 LDXSHNALQRLRRGWLAFLFQLRALHLDHNE LDALGRGVFVNASGLRLLDLSSNTLRALGRH DLDGLGALEKLLIFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHILTI.DL SSNRM  315 1665 A 2973 1 525 ITVSTHASGSPFGLEFQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTILDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT  316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVTGNVPKA GTDANVYLTYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYADLGALTKRRHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEB DDGGLSRE  317 1667 A 2981 3 440 VLNCQGRPTRFVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKTTM ATRKVIVTVQARPAVPMCWTEGQ RRTPGEGHVAAVDREVAGFPVPAEGISGETTH ENYDARLLRIDIANNILEQVGLIFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETTH	312	1662	A	2967	3	405	
313 1663 A 2969 2 430 VVADNCRQGYLDALRFLERRGLTKEPVLWT LVSKEPPAPADGNWDAGCDQRRKGGLSLNW KVPHVQVKDVPNFQLSPELBAALKKACTRD PSRWARFWHSGPGQVLTYLLLPCTLPFEYTYF RSRRLVVWLPDVPADLWWMQ  314 1664 A 2971 422 33 LDXSHNALQRLRPGWLAPLFQLRALHLDINE LDALGRGVFVNASGLRLDLSSNTLRALGRH DLDGLGALEKILLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM  315 1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQMEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT  316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTTYGEEPGDTGRPLKSSDKSNK FEQQQIDTFITYAIDLGALTKIRHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEB DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRNHGSLPHLQHRVRFAAS PSQYDASNILMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEPPIIRSSSLKVTKCLFTEQPKPHILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETTH						,	HTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV
DECEGPKPA							EAKLPWFKQAQELEEGAAVSEEPSFIPEAWS
1663							
LVSKEPPAPADGNWDAGCDQRRKGGLSLNW KVPHVQVKDVPNFEQLSPELBAALKKACTRD PSRWARFWHSGPGQVLTYLLLPCTLFFEYIYF RSRRLVWLPDVPADLWWMQ  314 1664 A 2971 422 33 LDXSHNALQRLRPGWLAPLFQLRALHILDHNE LDALGRGVFVNASGLRLLDLSSNTLRALGRH DLDGLGALEKLLIFNNRLVHLDEHAPHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM  315 1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGGVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVVIVAVLDLNDNT  316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTTYGEEYGDTGERPLKKSDKSNK FEQQQTDTFTTYADDLGALTKIRRHIDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGPPTRPVRINGDQQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRNVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ  318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETTH	212	1662	A	2000			
KVPHVQVKDVPNFEQLSPELĒAALKKACTRD   PSRWARFWHSGPGQVLTYLLLPCTLPFEYIYF   RSRRLVWULPDVPADLWWMQ     314   1664   A   2971   422   33   LDXSHNALQRLPGWLAFLFQLRALHLDHNE   LDALGRGVFVNASGLRILDLSSNTLRALGRH   DLDGLGALEKILLFNNRLVHLDEHAFHGLRA   LSHLYLGCNELASFSFDHLHGLSATHILTLDL   SSNRM     315   1665   A   2973   1   525   ITVSTHASGSPFGLEFQSGWLWVRAALDREA   QELYILKVMAVSGSKABLGQQTGTATVRVSI   LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV   FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ   TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP   RSTTGTVHVAVLDLNDNT     316   1666   A   2978   2   400   ELVVELVSAGKSGFERNTYEVQVVTGNVPKA   GTDANVYLTTYGEEYGDTGERPLKKSDKSNK   FEQGQTDTFTTYAIDLGALTKRIRHDNTGNR   AGWFLDRIDITDMNNEITYYFPCQRWLAVEE   DDGQLSRE     317   1667   A   2981   3   440   VLNCQGRPTRPVRINGDGQEVLYLAESDNVR   LGCPYVLDPDDYGFNGLDIEWMQVNSNPAH   HRENVFLSYQDKRINHGSLPHLQHRVFRAAS   DPSQYDASINLMNLQVSDTATYECRVKKTTM   ATRKVIVTVQARPAVPMCWTEGQ     318   1668   A   2995   119   414   LPEKEFPIRKSSSLKYTKCLFTEQPKPIILRFA   ENYDARILRIDIANTLREQVQELFNKTYGKQ   RRTPGEGHVAAVDREVAGFPVPAEGISGETTH	213	1003	A	2969	2	430	
PSRWARFWHSGPGQVLTYLLLPCTLPFEYIYF RSRRLVVWLPDVPADLWWMQ  314 1664 A 2971 422 33 LDXSHNALQRLRPGWLAPLFQLRALHLDHNE LDALGRGVFVNASGLRLLDLSSNTLRALGRH DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM  315 1665 A 2973 1 525 TIVSTHASGSPFGLEFQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRHPQ TGEVTILQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT  316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLITYGEEYGDTGERFLKKSDKSNK FEQQQTDTFTTYAIDLGALTKRIRHDNTGNR AGWFLDRIDTTDMNNEITYYFPCQRWLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKTTM ATRKVIVTVQARPAVPMCWTEGQ  318 1668 A 2995 119 414 LPEKEPPINKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETTH							
314 1664 A 2971 422 33 LDXSHNALQRLRPGWLAPLFQLRALHI.DHNE LDALGRGVFVNASGLRLLDLSSNTLRALGRH DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHIL.TILDL SSNRM  315 1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTILQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT  316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQIDTFITYAIDLGALTKIRRHDNTGNR AGWFLDRIDITDMNNEITYYFPCQR WLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ  318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRIPGEGHVAAVDREVAGFPVPAEGISGETIH					,		
1664							
LDALGRGVFVNASGLRLLDLSSNTLRALGRH DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM  1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTILQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT  316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTTYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYAIDLGALTKIRIHDNTGNR AGWFLDRIDTTDMNNEITYYFPCQRWLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH	314	1664	$\overline{\mathbf{A}}$	2971	422	33	
DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM  315  1665  A 2973  1 525  ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTILQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT  316  1666  A 2978  2 400  ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTTYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE  317  1667  A 2981  3 440  VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ BYDARLLRIDIANTLREQVGELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH	· · · · · · · · · · · · · · · · · · ·	-					LDALGRGVFVNASGLRLLDLSSNTLRALGRH
LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM  1 525 TTVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT  316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTIYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEPPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH			. [				
315 1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTILQTLDREQQSSYQLLVQVQDGGSPP RSITGTVHVAVLDLNDNT  316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTTYGEEYGDTGERPLKKSDKSNK FEQGQIDTFTTYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYPPCQRWLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ RTTPGEGHVAAVDREVAGFPVPAEGISGETIH	•	-	ſ	[	. [		
QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT  316  1666  A 2978  2 400  ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTTYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYAIDLGALTKRIRHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE  317  1667  A 2981  3 440  VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ TPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ RRTPGEGHVAAVDREVAGFPVPAEGISGETTH							COMPA.
LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT  316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTTYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ THE STORT OF THE STORT	315	1665	A T	2973	1	525	ITVSTHASGSPFGLEPQSGWLWVRAALDREA
LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT  316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTTYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ THE STORT OF THE STORT	I		·				QELYILKVMAVSGSKAELGQQTGTATVRVSI
TGEVTILQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT  316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEBYGDTGERPLKKSDKSNK FEQGQTDTFITYAIDLGALTKIRIRHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ  318 1668 A 2995 119 414 LPEKEPPIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETTH	1	i	ŀ	1	. [		LNQNEHSPRLSEDPTFLAVAENOPPGTSVGRV
RSTTGTVHVAVLDLNDNT  316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ  318 1668 A 2995 119 414 LPEKEPPIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETTH	l		}	1	}	}	FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ
316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTTYGEEYGDTGERPLKKSDKSNK FEQGQTDTFITYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ  318 1668 A 2995 119 414 LPEKEPPIRKSSSLKVTKCLFTEQPKPHIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETTH					l	Í	
GTDANVYLTTYGEEYGDTGERPLKKSDKSNK FEQGQTDTFITYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINILMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ ATRKVIVTVQARPAVPMCWTEGQ ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETTH	316	1666	<u>_</u>	2079	<del></del>	400	
FEQGQTDTFTTYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ  318 1668 A 2995 119 414 LPEKEPPIRKSSSLKVTKCLFTEQPKPHIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETTH	310 · }	1000	^	2710	-	+00	
AGWFLDRIDITDMNNEITTYFPCQRWLAVEE DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINILMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ  318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH		. [		ł	İ	1	
DDGQLSRE  317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINILMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ  318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH	l		ļ	ļ	l	1	AGMEL DEIDILDWINELLAALDVOOR VAGE
317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGFNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ  318 1668 A 2995 119 414 LPEKEFPIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH			ł	I	i		
LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ  318 1668 A 2995 119 414 LPEKEFPIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH	317	1667	A	2981	3	440	
HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ  1668 A 2995 119 414 LPEKEPPIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH							
DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ  318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH	{	ĺ	[	{	ſ	[	HRENVFLSYQDKRINHGSLPHLOHRVRFAAS
318 1668 A 2995 119 414 LPEKEPPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH			[	Ì	ſ	·	
ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH			l		]	J	ATRKVIVTVQARPAVPMCWTEGQ
ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH	318	1668	A	2995	119	414	LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA
RRTPGEGHVAAVDREVAGFPVPAEGISGETIH		i	ļ	ì	l		ENYDARLLRIDIANTLREQVQELFNKTYGKO
312 GFFAYTYGRLVVVEDLHSGAQQHWSGHSAEI	210	1660	<del>,  </del>	2000			RKIPGEGHVAAVDREVAGFPVPAEGISGETIH
	317	1003	A	<b>4999</b>	<u> </u>	332	GFFAYTYGRLVVVEDLHSGAQQHWSGHSAEI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	I Amino acid muse (A. Al. ). O. C. i.i.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		į	914	ng to first	acid residue	Q=Ghtamine, R=Arginine, S=Serine.
	l	l	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ľ	Ĭ			peptide		possible nucleotide deletion, possible
			!	sequence	l	nucleotide Insertion
						STLALSHSAQVLASASGRSSTTAHCQIRVWD
	· ·			İ		VSGGLCQHLIFPHSTTVLALAFSPDDRLLVTL
	ĺ.	l				GDHDGRTLALWGTGHL
320	1670	A	3000	693	322	IDESTGLIITVNYLDYETKTSYMMNVSATDQA
		1	1	]		PPFNQGFCSVYITLLNELDEAVQFSNASYEAA
	ĺ	ĺ		Ì	ĺ	ILENLALGTEIVRVQAYSIDNLNQITYRFDAY
						TSTQAKALFKIDAITVRGWGQGAPFFPI
321	1671	A	3001	6	383	RIPRGKACXTVLGRSTGELEGFASSRLPPQPC
		ĺ			•	GWGQSSDLLSRIDLDELMKKDEPPLDFPDTLE
l i		!	i :		ł	GFEYAFNEKGQLRHIKTGEPFVFNYREHLHR
						WNQKRYEALGEIITKYVYELLEKDCNSKKVS
322	1672	A	3007	192	447	ERVRNSLFPGRGDSQCACCPSSPVWVFLETGF
	i					LFPWLFLQVEVIKKAYMQGEVEFEDGENGK
323	1.000		2070			DGAASPRNVGHNIYILAHQLARH
323	1673	A	3019	18	245	KELLFYHLIVNNINFFNTRYAKIHIPIIASVSEH
	•					QPITWVSFFFDLHILVCTFPAGLWFCIKNIND
324	1674	A	3020	523	797	ERVFGKRGF
324	10/4	A	3020	523	197	LCYFSARYHQRKIFGILYIFTLSAINRKEPNLFI
				5.0	·	YLFIFFEMESHSVTHAGVQRHNLNSLQPLPPG
325	1675	A	3022	2	156	FKRFSCLCFLSSWNYRGAPPGPANF
322	1075	^	3022	2	136	NDFLPLYFGWVLTKKSSETLRKAGQVFLEEL
326	1676	Α	3023	38	172	GNHKAFKKELRQCRWQVGAL KMVRGSKKLISFFPGGPYGILAGRDPSKGLAT
520	10,0	^	3023	36	172	FCLNKEALKDEFE
327	1677	Ā	3027	1	385	LTLEFLLLPAASELAHGKRLACCIVDHKLPEC
		••	3027	•	363	GFYGLYDKILLFKHDPTSANLLOLVRSSGDIO
] }			•			EGDLVEVVLSASATFEDFQIRPHALTVHSYRA
1	ſ					PAFCDHCGEMLFGLVRQGLKCDGCGLNYHK
						RC
328	1678	A	3030	13	569	ITRPTISCQRPGPGLAAGMLPYTVNFKVSART
						LTGALNAHNKAAVDWGWQGLIAYGCHSLV
ĺ	ľ	- 1				VVIDSITAQTLQVLEKHKADVVKVKWAREN
l . <b>l</b>	ļ					YHHNIGSPYCLRLASADVNGKIIVWDVAAGV
						AQCEIQEHAKPIQDVQWLWNQDASRDLLLAI
						HPPNYIVLWNADTGTKLWKKSYADNILSFSF
300	1600					D
329	1679	A	3038	90	744	SVNLPPSLWPWEEAMDSTKSEPLKGSPEAED
ı	٠ ا					GNIEYKKLVNPSQYRFEHLVTQMKWRLQEG
l i				•		RGEAVYQIGVEDNGLLVGLAEEEMRASLKTL
1		1	ļ			HRMAEKVGADITVLREREVDYDSDMPRKITE
			1			VLVRKVPDNQQFLDLRVAVLGNVDSGKSTL
			i			LGVLTQGELDNGRGRARLNLFRHLHEIQSGR
			.			TSSISFEILGFNSKGEVHGINGTQWGQTLRMG
330	1680	A	3040	3	397	V
	.000	-7	2070	·	3)1	LCSTLLLLTIPSWVLSQITLKESGPTLMKPTET
			1		1	LTLTCTFSGFSLNTSGVGVAWIRQPPGKALE WLALIYWDDDKRYSPSLNDRLTIAKDTSRNO
.[	ŀ	. [	1		1	VVLTMTNMGPVDTATYYCAQFARGARGSN
		1	l		1	WFDPWGO
331	1681	A	3043	3	1509	AGIRHEAPPITSNRHRRQIDRGVTHLNISGLK
				-	1007	MPRGIAIDWVAGNVYWTDSGRDVIEVAQMK
ł		j	- 1	Ì		GENRKTLISGMIDEPHAIVVDPLRGTMYWSD
1		ł		l	l	WGNHPKIETAAMDGTLRETLVQDNIQWPTG
1	ł	ŀ	I	ŀ		LAVDYHNERLYWADAKLSVIGSIRLNGTDPI
i				-		VAADSKRGLSHPFSIDVFEDYIYGVTYINNRV
- 1	1		- 1	- 1	0	PKIHKFGHSPLVNLTGGLSHASDVVLYHQHK
l			Í	ŀ	ļ	QPEVTNPCDRKKCEWLCLLSPSGPVCTCPNG
ļ	J	1	J	}	J	KRLDNGTCVPVPSPTPPPDAPRPGTCNLQCFN
	_		l	1	İ	GGSCFLNARROPKCRCOPRYTGDKCELDOC
						7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  WEHCRNGGTCAASPSGMPTCRCPTGFTGPKC TQQVCAGYCANNSTCTVNGGNQPQCRCLPG
						FLGDRCQYRQCSGYCENFGTCQMAADGSRQ CRCTAYFEGSRCEVNKCSRCLEGACVVNKQS GDVTCNCTDGRVAPSCLTCVGHCSNGGSCT MNSKMMPECQCPPHMTGPRCEEHVFSQQQP GHIASILIP
332	1682	A	3045	3	952	TTTISNFHTQVNRTYCCGTYRAGPMRQISLVG AVDEEVGDYFPEFLDMLEESPFLKMILPWGT LSSLRLQCRSQSDDGPIMWVRPGEQMIPTAD MPKSPFKRRRSMNEIKNLQYLPRTSEPREVLF EDRTRAHADHVGQGFDWQSTAAVGVLKAV QFGEWSDQPRITKDVICFHAEDFIDVVQRLQ LDLHEPPVSQCVQWVDEAKLNQMRREGIRY ARIQLCDNDIYFIPRNVIHQFKTVSAVCSLAW HIRLKQYHPVVEATQNTESNSNMDCGLTGKR ELEVDSQCVRIKTESEEACTEIQLLTTASSSFP PASE
333	1683	A	3046	497	167	SACSTGPELPGRATRSLTRPANQKGCDGDRL YYDGCAMIAMNGSVFAQGSQFSLDDVEVLT ATLDLEDVRSYRAEISSRNLAVSAPVDTCVG CSSKTWKVAPFVRAWWRP
334	1684	A	3053	37	276	VITDLEEQLNQLTEDNAELNNQNFYLSKQLD EASGANDEIVQLRSEVDHLRREITEREMQLTS QKQVRRVNKVVRSLEDF
335	1685	A	3054	2	846	WDAWGDWSDCSRTCGGGASYSLRRCLTGR NCEGQNIRYKTCSNHDCPPDAEDFRAQQCSA YNDVQYQGHYYEWLPRYNDPAAPCALKCH AQGQNLVVELAPKVLDGTRCNTDSLDMCISG ICQAVGCDRQLGSNAKEDNCGVCAGDGSTC RLVRGQSKSHVSPEKREENVIAVPLGSRSVRI TVKGPAHLFIESKTLQGSKGEHSFNSPGVFVV ENTTVEFQRGSERQTFKIPGPLMADPIFKTRY TAAKDSVVQFFFYQPISHQWRQTDFFPCTVT CGGG
336	1686	A	3058	54	347	VVGKQEAGAHSDSCCLLHTPPRLTPAHSRKA LRNSRIVSQKDDVHVCIMCLRAIMNYQVSRG AWDWRLGSPACPHWGLHKLPRLWDPLSLYP VLCWGT
337	1687	A	3059	2	709	ILTSLVBLTRFETLTPRFSATVPPCWVEVQQE QQQRRHPQHLHQQHHGDAAQHTRTWKLQT DSNSWDEHVFELVLPKACMVGHVDFKFVLN SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF LEDHKEDILCGPVWLASGLDLSGHAGMLTLT SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI CNGGMRPVVRLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK
338	1688	A	3060	85	384	KAFYNYHVLELLQMLVTGGVSSQLEQHILDK DKVYGVADSCTSILLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP
339	1689,	A	3063	236	362	CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV
340	1690		3065	3	1249	DLWOFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL DNLGQTSLHRAAYCGHLQTCRLLLSYGCDPN

No. of No. of Deptide codde octation   No. of Deptide codde   No. of Deptide   No. of Dep	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
DISSN   Diestion   Display   Displ			hod			nucleotide	D=Aspartic Acid, E=Ghrtamic Acid,
Separation   Sep				-			F=Phenylalanine, G=Glycine, H=Histidine,
1692   1692   A 3073   463   3   1693   A 3075   250   1   11/14/APP/AMSALAGVSSVQVERAAQ   344   1694   A 3076   2   138   14/14/APP/AMSALAGVSSVQVERAAQ   345   1695   A 3078   469   3   14/14/APP/AMSALAGVSSVQVERAAQS   346   1696   A 3082   404   2   QNITSKDLDVSQVERIGINGE A SULGOVAN AGE AND AG			ľ			corresponding	I=Isoleucine, K=Lysine, L=Leucine,
manio said residue of peptide residue of peptide residue of sequence   T=Threonine, W=Valine, W=Typophun, T=Stop codon, peptide sequence   T=Threonine, W=Valine, W=Typoshibe nucleotide deletion, I=possible nucleotide deletion, I=possible nucleotide deletion, I=possible nucleotide insertion   ISI.OGFTALIOMGNINVO, ULDEGISLGNSER, DRQLLEARKAGDVETVEKLCTVSQSVNCRDI GRQSTPLIFFAAGYNRYSVVETLLQHGADVM AKDRGGL-YPLINACSYGHTEVALLLYCHGAVVAVVADLWKFTFLIEAAAKGKYEKLYLLQU GRQSTPLIFFAAGYNRYSVVETLLQHGADVM AKDRGGL-YPLINACSYGHTVEKLICTVSQSVNCRDI GRQSTPLIFFAAGYNRYSVETLLQHGADVM VAVVADLWKFTFLIEAAAKGKYEKLYLLQU GDAALLDAAKKGCLARVKKLSSPDNVNCRD TQGRISTPLIHAAGK   TQG		uence					M=Methionine, N=Asparagine, P=Proline,
Per	uence	]	ł	914			Q=Glutamine, R=Arginine, S=Serine,
Popsible nucleotide deletion,   popsible nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide   nucleotide insertion   nucleotide   nuc	1						V-Typesine V-II-lenen * Oten and a
	}		J			Soquetice	======================================
	ľ		1			1	nucleotide insertion
DRQLLEAAKAGDUFUNTUSVISTLÜGHADVH		<del></del>		<del>                                     </del>			
GRQSTFLIFAAGYNRVSVYSYLLÓHAGMA	l						DROLLEAAKAGDVETVKKLCTVOSVNCRDIE
AKDRGGLYPLHNACSYGHYEVAELLVKHAGA				1	ĺ	Í	GROSTPLHFAAGYNRVSVVEYLLOHGADVH
1691   A 3070   1 547   GVLIPSFQNQLFADILAGIESVTSEINYQTILAG   NYNYDARDSEESVINLLSYNDOGIIDEQUILI GDAALLDAAKKGCLAAVKKLSSPDNVNCRD TQGHSTPLILKAGK   TQGHSTPLILKAGK   TQGHSTPLILKAGK   TQGHSTPLILKAGK   TQGHSTPLILKAGK   TQGHSTPLIL				1			AKDKGGLVPLHNACSYGHYEVAELLVKHGA
GBAALDAAKGGCLARVKKLSSPDNVNCRD			Ì	•			VVNVADLWKFTPLHEAAAKGKYEICKLLLQ
1691	ł	ł	ł	ł	ł	}	HGADPTKKNRDGNTPLDLVKDGDTDIQDLLR
1691		•		ł	[		
NYNYDROSEESVINLLSYNIDGIILSEKYHTI   RTVK4RRSATIPVVELMDVQGERLDMGVGET    NRQAAFDMVCYMLEKKYRHKILVLGSKDDT    RDEQRYQGYCDAMMLHNLSPLRMPRAISS    HLRMQLMRDALSANPDLDGYCTN    RDEQRYQGYCDAMMLHNLSPLRMPRAISS    HLRMQLMRDALSANPDLDGYCTN    RDEQRYQGYCDAMMLHNLSPLRMPRAISS    HLRMQLMRDALSANPDLDGYCTN    RDEQRYQGYCDAMMLHNLSPLRMPRAISS    HLRMQLMRDALSANPDLDGYCTN    A 3073   463   3   RINRCKFSDADILVPGPTISLIGITISLIGIPNI     DEDRAVTAGEVDLLRGGEKAPVMAATRIRL     A YSGVRPLVASDDDPSGRNVSRGIVLLDHAB     RDGLDGFTITTGGKLMTYMAAPWATDAVC     RRLGMTRCTTADLALPGSQEPAKVP     RRLGMTRCTTADLALPGSQEPAKVP     RRLGMTRCTTADLALPGSQEPAKVP     RRLGMTRCTTADLALPGSQEPAKVP     STLVAABLAATRGLGFM     STLVAABLAATRGLAAFS     SLLGCDVQDADDFTALGGHISLAAFS     SLLGCDVQDADDFTA	241	1601	<u> </u>	2070		646	
	341	1091	A	3070	ļ <b>1</b>	347	GVLIPSFQNQLFADILAGIESVTSEHNYQTLIA
NRQAAPDMVCTMLEKRYRHKIL/LGSKDDT	İ						
RDEORYOGYCDAMM.HNI.SPLRMMPRAISS   HLRMQi.mRADLSAMPDLDGYPCTN							
HLRMQLMRDALSANPDLDGYFCTN		}		]	]		
342			ł	Ì			
DIDNRYTAEEVPOLLREGEKLAPVMAKTRIL   AYSGVRPLVASDIDPGRNVSGGIVLIDHAE   ROGLOGFITTIGGKLMTYRLMAEWATDAVC   RKLGNTRPCTTADLALPGSGEPAKVP	342	1692	A	3073	463	3	
AYSGYRPLVASDDDPSGRNVSRGIVLIDHAE   RDGLDGFITTIGGKMTYRLMAEWATDAVC   RKLGNTRPCTTADLALPGSQEPAKVP   RKLGNTRPCTTADLALPGSQEPAKVP   RKLGNTRPCTTADLALPGSQEPAKVP   RKLGNTRPCTTADLALPGSQEPAKVP   ULTYLAIFAPVAMSALAGVKSVQQVRIRAAQS   LGASRAQVLWPVJGALPGELTGEIGLGGVG   WSTLVAABLIAATRGLGFM   STANDALIAATR		1	ļ				IDDNRVTAEEVDILLREGEKLAPVMAKTRILR
RKLGNTRPCTTADLALPGSQEPAKYP	ļ	İ	<u> </u>				AYSGVRPLVASDDDPSGRNVSRGIVLLDHAE
1693   A   3075   250   1		ľ	ł	1			
1694   A   3076   2   138	2/2	1600	Ļ.—	2000			
WSTLVAAELIAATRGLGFM	343	1693	A	3075	250	1	LLIYLAIFAPVAMSALAGVKSVQQVRIRAAQS
344		]					LGASRAQVLWFVILPGALPEILTGLRIGLGVG
345	3/1/	1604	A	2076	2	120	
1695	]	1074	Α.	3070	2	130	
VINQAAATGGDARQLVGYILVSQGLPLDTSA	345	1695	A	3078	469	3	
LQAQLRETLPPHMVPVVLLQLPQLPLIANGKI							
DRKALPLPELKAQAPGRAPKAGSETILAAAFS   SLLGCDVQDADADFFALGGHSLLAMKLAT   SLLGCDVQDADADFFALGGHSLLAMKLAT   QNTSKDLDVRLDPQTVPIELEQL.VLSPNHMI   ERIEDVFTRQSNFSADIAHEIRTPITNLITQTEI   ALSQSRSQKELEDVLYSNLEELTRMAKMVSD   MLFLAQADNNQLIPEKKMLNLAHEVGKVFD   QFEALPE	i						LOAOLRETLPPHMVPVVLLOLPOLPLIANGKI
SLLGCDVQDADADFFALGGHSLLAMKLAT						. ,	DRKALPLPELKAQAPGRAPKAGSETIIAAAFS
BRIEDVFTRQSNFSADIAHEIRTPTINLITQTEI ALSQSRSQKELEDVLYSNLEELTRMAKMVSD MLFLAQADNNQLIPEKKMLNLAHEVGKVFD QFEALPE  347 1697 A 3084 3 340 NELTFKEAEISKLYTKVHPAYRTLLEKRQALE DEKAKLNGRVTAMPKTQQEIVRLTRDVESGQ QVYMQLLINKEQELKITEASTVGDVRIVDPAIT QPGVLKPKKGLIILGAI QPGVLKPKKGLIILGAI QPGVLKPKKGLIILGAI  348 1698 A 3086 723 10 TQAMVWQQKACAEDDPQLSGRHWLHAATL YNIAAYPHLKGDDLAEQAQALSNRAYEEAA QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFARGI AMLTIDMPSVGFSSK WKLTQDSSLLHQHVLK ALFNVPWVDHTRVAAFGFRFGANVAVRLAY LESFRLKAVACLGPVVHTILLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH ALFNVPWVDHTRVAAFGFSLGAQVCG SSLWIVLAAVGIGAWNM  349 1699 A 3087 2 249 RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIVLAAVVGIGAWNM  350 1700 A 3099 3 424 EAPEATPQPSQFGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVYMVETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREQLFMANYGADDWR							SLLGCDVQDADADFFALGGHSLLAMKLAT
ALSQSRSQKELEDVLYSNLEELTRMAKMVSD MLFLAQADNNQLIPEKKMLNLAHEVGKVFD QFEALPE  347 1697 A 3084 3 340 NELTFKEAEISKLYTKVHPAYRTLLEKRQALE DEKAKLNGRVTAMPKTQQEIVRLTRDVESGQ QVYMQLLINKEQELKITEASTVGDVRIVDPAIT QPGVLKPKKGLIILGAI  348 1698 A 3086 723 10 TQAMVWQQKACAEDDPQLSGRHWLHAATL YNIAAYPHLKGDDLAEQAQALSNRAYEEAA QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFARAGA AMLTIDMPSVGFSSK WKLTQDSSLLHQHVLK ALPNVPWVDHTRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTILLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH 349 1699 A 3087 2 249 RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWWLAAVGIGAWNM  350 1700 A 3099 3 424 EAPEATPQPSQFGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIVVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREQLFMANYGADDWR	346	1696	A	3082	404	2	QNITSKDLDVRLDPQTVPIELEQLVLSFNHMI
MLFLAQADNNQLIPEKKMLNLAHEVGKVFD QFEALPE  347 1697 A 3084 3 340 NELTFKEAEISKLYTKVHPAYRTLLEKRQALE DEKAKLNGRVTAMPKTQQEIVRLTRDVESGQ QVYMQLLNKEQELKITEASTVGDVRIVDPAIT QPGVLKPKKGLIILGAI  348 1698 A 3086 723 10 TQAMVWQQKACAEDDPQLSGRHWLHAATL YNIAAYPHLKGDDLAEQAQALSNRAYEEAA QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI AMLTIDMPSVGFSSK WKLTQDSSLLHQHVLK ALPNVPWVDHTIRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH  349 1699 A 3087 2 249 RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM  350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR							ERIEDVFTRQSNFSADIAHEIRTPITNLITQTEI
QFEALPE							
1697   A   3084   3   340   NELTFKEAEISKLYTKVHPAYRTLLEKRQALE   DEKAKLNGRVTAMPKTQQEIVRLTRDVESGQ QVYMQLLNKEQELKITEASTVGDVRIVDPAIT QPGVLKPKKGLIILGAI     348   1698   A   3086   723   10   TQAMVWQQKACAEDDPQLSGRHWLHAATL YNIAAYPHLKGDDLAEQAQALSNRAYEEAA QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI AMLTIDMPSVGFSSK WKLTQDSSLHQHVLKA ALPNVPWVDHTRVAAFGFFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQVVPE MYLDVLASRLGMHDASTKSSTRENH     349   1699   A   3087   2   249   RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM     350   1700   A   3099   3   424   EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK     351   1701   A   3108   2   404   IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR						·	
DEKAKLNGRVTAMPKTQQEIVRLTRDVESGQ QVYMQLLNKEQELKITEASTVGDVRIVDPAIT QPGVLKPKKGLIILGAI  348 1698 A 3086 723 10 TQAMVWQQKACAEDDPQLSGRHWLHAATL YNIAAYPHLKGDDLAEQAQALSNRAYEEAA QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI AMLTIDMPSVGFSSK WKLTQDSSLLHQHVLK ALPNVPWVDHTRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH  349 1699 A 3087 2 249 RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM  350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSRDVVSPNCSNTVQEKTFN KDTVIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR	347	1697	A	3084	3	340	
348 1698 A 3086 723 10 TQAMVWQKACAEDDPQLSGRHWLHAATL YNIAAYPHLKGDDLAEQAQALSNRAYEEAA QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI AMLTIDMPSVGFSSKWKLTQDSSLLHQHVLK ALPNVPWVDHTRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH 349 1699 A 3087 2 249 RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM 350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSBDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK 351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR	3.,,	1000		3007		340	
348 1698 A 3086 723 10 TQAMVWQKACAEDDPQLSGRHWLHAATIL YNIAAYPHLKGDDLAEQAQALSNRAYEEAA QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI AMLTIDMPSVGFSSK WKLTQDSSLLHQHVUK ALPNVPWVDHTRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQVPE MYLDVLASRLGMHDASTKSSTRENH  349 1699 A 3087 2 249 RIRSSDPEITLAGTPLHAAYLIGMTILCAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGA WNM  350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR							
348 1698 A 3086 723 10 TQAMVWQQKACAEDDPQLSGRHWLHAATL YNIAAYPHLKGDDLAEQAQALSNRAYEEAA QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI AMLTIDMPSVGFSSK WKLTQDSSLLHQHVUK ALPNVPWVDHTRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQVPE MYLDVLASRLGMHDASTKSSTRENH  349 1699 A 3087 2 249 RIRSSDPEITLAGTPLHAAYLIGMITLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM  350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR						·	
YNIAAYPHI.KGDDLAEQAQALSNRAYEEAA QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI AMLTIDMPSVGFSSK WKLTQDSSLHQHVLK ALPNVPWVDHTRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH  349 1699 A 3087 2 249 RIRSSDPEITLAGTPLHAAYLIGMITLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM  350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR	348	1698	A	3086	723	10	
QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI AMLTIDMPSVGFSSK WKLTQDSSLLHQHVLK ALPNVPWVDHTRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH  349 1699 A 3087 2 249 RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGA WNM  350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR	•						
AMLTIDMPSVGFSSKWKLTQDSSLLHQHVLK ALPNVPWVDHTRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH  349 1699 A 3087 2 249 RIRSSDPETTLAGTPLHAAYLIGMTLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM  350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR		l i		1	i	ŀ	QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG
ALPNVPWVDHTRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH  349 1699 A 3087 2 249 RIRSSDPETTLAGTPLHAAYLIGMTLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM  350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR							PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI
LESPRLKAVACLGPVVHTLLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH  349 1699 A 3087 2 249 RIRSSDPEITLAGTPLHAAYLIGMITLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM  350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR					·		
349 1699 A 3087 2 249 RIRSSDPEITLAGTPLHAAYLIGMILICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM 350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYPK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK 351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR							
349 1699 A 3087 2 249 RIRSSDPEITLAGTPLHAAYLIGMITLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM 350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK 351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR						0	
GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM  350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR	349	1690	Δ	3087	2	240	
350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK 351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR	5-17	(10)	^	3007	-	247	
350 1700 A 3099 3 424 EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR					ĺ		
ANTPOSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR	350	1700	A	3099	3	424	
NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK  351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR		_			-		
KDTVIIVSEPSEDEESQGLPTMARRNDDISELE   DLSGMEDLK     351   1701   A   3108   2   404   IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG   MFGIVVMVIETELSWGAYYKAPLYSLALKCL   ISLFTIILLGLTIVYHAREIQLFMANYGADDWR					j		·
351 1701 A 3108 2 404 IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR						i	
MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR				<b> </b>			
ISLFTIILLGLTIVYHAREIQLFMANYGADDWR	351	1701	A	3108	2	404	IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG
			Ì				
			- 1	1			ISLFTIILLGLTIVYHAREIQLFMANYGADDWR
SALTYEPIFLILLEALRGVIHATPCRVSLSLWD		J	J		ļ		
GLDLP		L					GLDLP

SEQ ID	SEQ ID	1 1/-4	Largo	15 0		
NO: of	NO: of	Met hod	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	noa	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-	1	in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence	1	09/496	correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine,
uence	duice	ł	914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline,
1.01.00			714	amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine,
	1		ĺ	residue of		T=Threonine, V=Valine, W=Tryptophan,
	!	1		peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				sequence		/=possible nucleotide deletion, \=possible
352	1702	A	3110	341	2	nucleotide insertion
552	1,02	1	3110	341	4	AQLAEVCPPQTLLTTNTSSISITAIAAEIKNPER
1	1	l	1			VAGLHFFNPAPVMKLVEVVSGLATAAEVVE
1	1		1			QLCELTLSWGKQPVRCHSTPGFIVNRVARPY
353	1703	A	3111	3	188	YSEAWRALEEQVAAPEVI
	1705		]	3	100	HFSLFRIAFAVFLTYMTVGLPLPVIPLFVHHEL
354	1704	A	3116	367	225	GYGNTMVGIAVQIQFLATVLTRGYAGRLA
) 331	1704	] ^	3110	307	223	WQLFHLNGTFLNIGETDTESCVNGWVYDRSS
355	1705	A	3117	101	53	FPFSNMTEVRGLVFLS
1 333	1,03	Δ	3117	101	33	VINLVYLISSPRPELKPVDKESEVVMKFPDGF
1	ļ		1			EKFSPPILQLDEVDFYYDPKHVIFSRLSVSADL
		ľ	ſ			ESRICVVGENGAGKSTMLKLLLGDLAPVRGI
			ļ			RHAHRNLKIGYFSQHHVGAAGT*TFSACGNL
1.						LGTQVFLGRPEEEY\RHQLGFGMGISGELGHA
ſ	ĺ .	ĺ	ľ			SSLPACLGGQKEAEVAFCSDGLLPCPNFL\IL\
1						DEPTN/HLGHGRAIEALGPCLQTISGVGVILVS
356	1706	A	3121	137	466	HE*SALSRLVCRELWVC*GRSTSPF
1	1	1.	3121	137	400	RGGRDWGEHNQRLEEHQARAWQGAMDAG
Į.	ļ	J	]			AASREHARWQGTGLAPGTRVAVAPTCVQGL
j						POERSVCRPFFSSRWREGPVWALGAGAHGKP
357	1707	A	3124	1249	229	RWSGGVRCVVRGGRWFTPAPH
1 22.	] -//	^	3124	1249	229	MLEAPGPSDGCELSNPSASRVSCAGQMLEVQ
1			[ ]	·		PGLYFGGAAAVAEPDHLREAGITAVLTVDSE EPSFKAGPGVEDLWRLFVPALDKPETDLLSH
İ				-		LDRCVAFIGQARAEGRAVLVHCHAGVSRSV
	ا ،		i i			AIITAFLMKTDQLPFEKAYEKLQILKPEAKMN
1			[ [	ĺ		EGFEWQLKLYQAMGYEVDTSSAIYKQYRLQ
ŀ			1		İ	KVTEKYPELQNLPQELFAVDPTIVSQGLKDE
						VLYKCRKCRRSLFRSSSILDHREGSGPIAFAH
{	1		1 1	1		KRMTPSSMLTTGRQAQCTSYFIEPVQWMESA
1	٠	•		J		LLGVMDGQLLCPKCSAKLGSFNWYGEQCSC
			]			GRWITPAFQIHKNRVDEMKILPVLGSQTGKI
358	1708	Α	3127	816	139	EVETLGPRTPGP/EAQSPTPGSCPGWQEPSPGP
				J		TPPP*LSGPGPQGAPVLGKLLPDPEETPAGKTP
						LGKHFWWGL\PVTSANFSPGAAA*FGGALSPP
						GGDL/GHMLLQGPPSPFRLQQQ*QTPPGSHSP
,	j			1		PTANREINPOPAAAADTRSCWGHKRSWRGW
i i			l l	ĺ	ſ	RGLAPWRLGFGSPGIP*PAPAGIP/GRPTWEGG
						KGAGGKPSETLTRSPPVWRGKRGSANGFLSW
						VQILQ
359	1709	Α	3132	3	191	HEHLLLLLCVFLVKSQGVNDNEEGFFSARG
'				j	ļ	HRPLDKKREDAPNLRPALAD\ITVCDYRAQIA
						*AASTPKRAASIAHNAVSCR*AQIA
360	1710	A	3134	1	286	REPPRPALLFF*DRVSLCCPGWNAVVQSQLT
	. 1			ļ	j	AAPTSQVQ/SDSPTFPSSWDYRHVPEYPANFL
	ŀ					*RQGFPMLPRLVSNSWAQTVHPPRPPKVLDL
					<u> </u>	QA
361	1711	A	3135	56	1449	PVPAPRVSPSARGAPGRPRLPGVRGPRHS/WA
٠	Ì		·		[	AD*RGSRM/PPRAPAPSPTGP/APGGKKVRGR
		l		1	l	VPEDPDAYEPRCSAL*V*PTHVTSPQFCDP*N
	I				į	GQIRSYFTVLLRGLNETMLVK/PLCRREP/PEA
	i				į	GPGRQSTPAVTRDHRQHEDPRGAGRQWDAD
					l	PRPSAP/PAEVATGSRPGRHMWMRLCLAAQQ
	ļ	- 1	i	. ]	J	APGLPHRTSIRPGWRRLTEPEAWARRHRRPW
		ł		ļ		GQRGAVRPPPQGAAPPPSHQGRRTNTDPSAT
	1	ł	ł	1		PRLTVMSRCLAPDLKAPASGPRGWRRGMPQ
ĺ			. 1	ļ	.]	SS/GALLWTPPPTPRGSHSPRPREAPLRAIHPA
		J	·	. 1	1	GPSK/SRAGASGRLPEVIYGWVTLFTPPEAGT
·	ł	1	ŀ	1		F/LIPSPT*MSPALVIQPPVPPTOMGLRISGLPR
						QG*PSGAPW*LPGLAQLAFQCHLPHDEVGPP

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cystcine, D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		١.	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
i i	l	ľ	ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	·		1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
<u> </u>				sequence		nucleotide insertion
ļ	}	l				RNQSPLGNDTLSSGLPMGPRRQVWPLARVG
1						GHSSPREPQVLKKPLWGQTDIAGVGSASLYP
362	1712	A	3136	1050		DNL
302	1/12	A	3130	1270	274	RVGMVLGTREVGDSTPPPSPPLYPFTGNEFVQ
		l				HNTWQLSRVYPSDLRTDSSNYNPQELWNAG
1 .	ļ	]				CQM/V*GGSRDWEEGVEEQQVGNKFSSDGR
1		l .	İ			VGECSRKLLG*EMLSVDITSRYRAPSTYLLNS LKEGLEGLHGESCSSFLLGPSVAMNMQTAGL
						EMDICDGHFRQNGGCGYVLKPDFLRDIQSSF
						HPEKPISPFKAQTLLNQVISVQQLPKVDKTKE
	}	1	ł			GSIVDPLVKVQIFGVRLDTARQETNYVENNG
						FNPYWGQTLCFRVLGPDFPMLRFGKMDYDW
						KSRNDLLGKTPCPGTCMQQGYRHIHLLSKDG
				1	·	ISLRPASIFVYICIQEGLEGDES
363	1713	С	3139	60	248	MFAGSYGKSMFSFSKKVLNCLPKWRYHFVIA
						PAMNESPLAPHLHQHLVFSVFQVLTILIGV**
364	1714	A	3140	57	418	SAFKTLQLPAFSLYFDLGSLKLLILRIHTSIVK
1						NHKVESPRTMSPG*DPQSFLQIPQPRPPQLRV
						GLTSGLIQHFHSPSSCQFPLLRGPPFPRQPPLGI
365	1015					SGASLCPVLSPPR*PLQPSSL
365	1715	A	3145	122	413	LLPYPSLFVFLRQCHFVTRLECNGVVSAHCN
i I						LHLPGSSDSPASAS*VAGTTGVCHHTRLIF\VF
] ]						LV*TGFHYVAQAGLELLTA*S\PPQLPKVVGL
366	1716	A	3150	247	2	QA
300	1710	^	3130	247		VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYY/NREPQNGRKIFANYVS\DK
i						GLMATIYEELLKLSNKLIQ
367	1717	A	3152	3	2367	QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT
				_		PIEKSDFAAATHPRAFYLSKPDETPNAWMSD
				İ		SGTGLTYWKLEEKDMHHSLPETLEKTFISLSS
		- 1			· j	TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP
1			ļ	İ		NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN
ì		- 1				VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE
			}			SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS
		ł	]			ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR
i 1		j				QNLKEKHARHIADLRAYYESEINSLKQKLEA
	ſ	•				KEISGVEDWKITNQILVDRCGQLDSALHEATS
		1				RVRTLENKNNLLEIEVNDLRERFSAASSASKI
	1	ŀ			,	LQERIEEMRTSSKEKDNTIIRLKSRLQDLEBAF
		1				ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR
			l	1		MISKLEAQVKQVEHENMLSLRHNSRIHVRPS
	]			.	•	RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT
}	. 1	ł		l		QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST
						SLLIKKQRETSDTPIMRALKELDEGKIFKNWG
i	j	ļ	]	]		TOTEKEDTSNSLL*/INPROTETSVNASRSPEK
		ļ	ĺ	1		CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR
	Ì	1	·			EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN
	ļ	j			İ	ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV
	l	ļ	ļ	ļ	j	KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN
.	,		}	j		DFEYTAKIRTLAETERFFDELTKEKDQIEAAL
260	1010					SRMPSPGGRITLQTRLNQVKCLSLNLL
368	1718	A	3163	2	2350	EFKSGGCGAGLVAAGAVLVLYPASRAGERT
		1	. 1	l		RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL
	j			l	. [	LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ
1		l		ļ	• 1	MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT
ł		İ		ţ		TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK
-		İ		İ		NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP
	<u> </u>					AWDLKGQLCDLNAELKRCRERTQTLDQENQ

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, I=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
	·					QLQDQLRDAQQQVKALGTERTTLEGHLAKV QAQAEQGQQELKNLRACVLELEERLSTQEGL VQELQKKQVELQEERRGLMSQLEEKERRLQT SEAALSSSQAEVASLRQETVAQAALLTEREER LHGLEMERRRLHNQLQELKGNIRVFCRVRPV LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLSRSD ERRGTLSGAPAPPTRHDFSFDRVFPPGSGQDE VFEEIAMLVQSALDGYPVCIFAYGQTGSGKTF TMEGGPGGDPQLEGLIPRALRHLFSVAQELSG QGWTYSFVASYVEIYNETVRDLLATGTRKGQ GGECEIRRAGPGSEELTVTNARYVPVSCEKEV DALLHLARQNRAVARTAQNERSSRSHSVFQL QISGEHSSRGLQCGAPLSLVDLAGSERLDPGL ALGPGERERLRETQAINSSLSTLGLVIMALSN KESHVPYRNSKLTYLLQNSLGGSAKMLMFV NISPLEENVSESLNSLRFASKVEPSVLFGTAQS NRKWKTDPDLCVCVCVCVCVCVCVCVCVP MSMYRVRGGRVAGGCFIGWRAPCPRAIK
369	1719	A	3165	365	12	GYTSQGRWIDIERGPLTANTESLHENNFNALP GYIRKIE*I*IYKKN*INFGGVGLLNIVKISILS/K IYRFDAIPVKILTRFFINLDKLILKFVLKTKIAK NRIKTFYIMRRKKLGDSS
370	1720	Α.	3170	393	42	GASISPSAVIDGVEGLKPMQEQEAQEAGPCLD *HMAPEQWVAPR\RLLFRLIFSVLHALIIAAAA QSSAEEDEDPRN*GQSSEDQAPNQNGLIVIVH RVHVPLGAAATVPVHRSHFPR
371	1721	A	3173	770	510	GNGGCGLSQIPPSHLGAFSRGSLLSRG\DPRGP PPHPVIFFVFVVE\QGFTVLARMVSIS*PCDPP ALASQSAGITGVSHLARPQNLYF
372	1722	·A	3180	381	76	RVLHHDNVPAHSSPQKREISQEFQLEIRHLP*S PDLAPSGCFLFLNLKNIFK\GTHFSLVDNVKK TVSTWLH/SQNAQFYKDRLNGWYHCLQKCL QHY*AYVEK
373	1723	A	3181	410	14101	RREVAGPEGKGLLLASAHTMLTPPLLLLPLL SALVAAAIDAPKTCSPKQFACRDQITCISKGW RCDGERDCPDGSDEAPEICPQSKAQRCQPNE HNCLGTELCVPMSRLCNGVQDCMDGSDEGP HCRELQGNCSRLGCQHHCVPTLDGPTCYCNS SFQLQADGKTCKDFDECSVYGTCSQLCTNTD GSFICGCVEGYLLQPDNRSCKAKNEPVDRPP VLLIANSQNILATYLSGAQVSTITPTSTRQTTA MDFSYANETVCWVHVGDSAAQTQLKCARM PGLKGFVDEHTINISLSLHHVEQMAIDWLTGN FYFVDDIDDRIFVCNRNGDTCVTLLDLELYNP KGIALDPAMGKVFFTDYGQIPKVERCDMDG QNRTKLVDSKIVFPHGITLDLVSRLVYWADA YLDYIEVVDYEGKGRQTIIQGILIEHLYGLTVF ENYLYATNSDNANAQQKTSVIRVNRFNSTEY QVVTRVDKGGALHIYHQRQPRVRSHACEN DQYGKPGGCSDICLLANSHKARTCRCRSGFS LGSDGKSCKKPEHELFLVYGKGRPGIIRGMD MGAKVPDEHMIPIENLMNPRALDFHAETGFI YFADTTSYLIGRQKIDGTERETILKDGIHNVE GVAVDWMGDNLYWTDDGPKKTISVARLEK AAQTRKTLIEGKMTHPRAIVVDPLNGWMYW TDWEEDPKDSRRGRLERAWMDGSHRDIFVT SKTVLWPNGLSLDIPAGRLYWVDAFYDRIETI LLNGTDRKIVYEGPELNHAFGLCHHGNYLFW TEYRSGSVYRLERGVGGAPPTVILLRSERPPI

	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
	nucl-	peptide	]	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
	eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
	uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Scrine,
		ĺ		714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
					residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				<b>j</b>	peptide	sequence	i-lyrosine, A=Unknown, =Stop codon,
					sequence		/-possible nucleotide deletion, \-possible
					sequence		nucleotide insertion
						·	FEIR/MYDAQHQQVGSNKCRVNNAGCSSLCL
							ATPGSRQCACAEDQVLDADGVTCLANPSYVP
							PPQCQPGEFACANSRCIQERWKCDGDNDCLD
							NSDEAPALCHQHTCPSDRFKCENNRCIPNRW
							LCDGDNDCGNSEDESNATCSARTCPPNQFSC
							ASGRCIPISWTCDLDDDCGDRSDESASCAYPT
							CFPLTQFTCNNGRCININWRCDNDNDCGDNS
							DEAGCSHSCSSTQFKCNSGRCIPEHWTCDGD
							NDCGDYSDETHANCTNQATRPPGGCHTDEF
Ų						,	QCRLDGLCIPLRWRCDGDTDCMDSSDEKSCE
- 1							GVTHVCDPSVKFGCKDSARCISKAWVCDGD
ı			·	-			NDCEDNSDEENCESLACRPPSHPCANNTSVC
Į				]	j	-*-	LPPDKLCDGNDDCGDGSDEGELCDQCSLNN
			l		ł	·	GGCSHNCSVAPGEGIVCSCPLGMELGPDNHT
						i	CQIQSYCAKHLKCSQKCDQNKFSVKCSCYEG
-					ļ		WVLEPDGESCRSLDPFKPFIIFSNRHEIRRIDLH
- [		1			1		KGDYSVLVPGLRNTIALDFHLSQSALYWTDV
١	1	i	.				VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG
Į	- 1	1	- 1	]			LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT
-1				i		·	TLLAGDIEHPRAIALDPRDGILFWTDWDASLP
١		]	1		ļ		RIEAASMSGAGRRTVHRETGSGGWPNGLTV
ı	}	J	J		j		DYLEKRILWIDARSDAIYSARYDGSGHMEVL
ļ	i						RGHEFLSHPFAVTLYGGEVYWTDWRTNTLA
ı	•					1	KANKWTGHNVTVVQRTNTQPFDLQVYHPSR
- 1	1	i	, i	ļ	)		QPMAPNPCEANGGQGPCSHLCLINYNRTVSC
-	ŀ			ļ			ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR
١	1	ł					GVDLDAPYYNYIISFTVPDIDNVTVLDYDARE
1	l	ł	- 1	1	-		QRVYWSDVRTQAIKRAFINGTGVETVVSADL
1		i			İ	•	PNAHGLAVDWVSRNLFWTSYDTNKKQINVA
J							RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY
ł	ł	- 1	- 1	1	ł	1	WTDGDNISMANMDGSNRTLLFSGQKGPVGL
1	l	ŀ	- 1			ľ	AJDFPESKLYWISSGNHTINRCNLDGSGLEVID
1					f		AMRSQLGKATALAIMGDKLWWADQVSEKM
ł	ł		- 1	ł	ł	ì	GTCSKADGSGSVVLRNSTTLVMHMKVYDESI
1	l	ļ	ł	i		İ	QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC
	- 1	İ	Ì	1			MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR
1	. 1	1	ľ		ł	1	GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND
1						ļ	TTYWVDMGLSTISRAKRDQTWREDVVTNGIG
	1	1					RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG
1	f	ı	ŀ	1	ł	· 1	SFRYVVISQGLDKPRAITVHPEKGYLFWTEW
1		l	l			j	GQYPRIERSRLDGTERVVLVNVSISWPNGISV
1				- 1	Į.	1	DYQDGKLYWCDARTDKIERIDLETGENREVV
1	ł	ľ	Į			1	LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK
ı	.		- [		l		RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR
	l	ļ.		1			QKGTNVCAVANGGCQQLCLYRGRGQRACA
1	-	- 1	· j	- 1	t	ł	CAHGMLAEDGASCREYAGYLLYSERTILKSI
	.		İ		ŀ		HLSDERNLNAPVQPFEDPEHMKNVIALAFDY
	ļ	.	1	0.0	ľ		RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT
1	1	1	1	ł	ŀ	ł	IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT
1	ł	ļ			1		RHTVDQTRPGAFERETVITMSGDDHPRAFVL
1		1	1		ŀ		DECQNLMFWTNWNEQHPSIMRAALSGANVL
1	1	ĺ	- 1	- 1	- 1	İ	TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE
			ł		1		RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF
1	-		- 1		l	į	WTDWVRRAVQRANKHVGSNMKLLRVDIPQ
ı	İ	1	- (	ľ	}	·	QPMGIIAVANDTNSCELSPCRINNGGCQDLCL
I	.		1	1	1		LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR
1		].		i	1	1	AQDEFECANGECINFSLTCDGVPHCKDKSDE
1	f		- 1	. 1	ĺ	Ī	KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN
	- 1	1	1		1	1	GADDCGDG\$DEIPCNKTACGVGEFRCRDGTC
L							IGNSSRCNQFVDCEDASDEMNCSATDCSSYF

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence .		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		}		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ľ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		i I		peptide		/=possible nucleotide deletion, \=possible
<u> </u>	<del> </del> -	ļ		sequence		nucleotide insertion
						RLGVKGVLFQPCERTSLCYAPSWVCDGAND
						CGDYSDERDCPGVKRPRCPLNYFACPSGRCIP MSWTCDKEDDCEHGEDETHCNKFCSEAQFE
	}					CONHRCISKOWLCDGSDDCGDGSDEAAHCE
		l				GKTCGPSSFSCPGTHVCVPERWLCDGDKDCA
						DGADESIAAGCLYNSTCDDREFMCONROCIP
			'			KHFVCDHDRDCADGSDESPECEYPTCGPSEF
						RCANGRCLSSRQWECDGENDCHDQSDEAPK
1		l				NPHCTSPEHKCNASSQFLCSSGRCVAEALLCN
1						GQDDCGDSSDERGCHINECLSRKLSGCSQDC
	1					EDLKIGFKCRCRPGFRLKDDGRTCADVDECS
						TTFPCSQRCINTHGSYKCLCVEGYAPRGGDP
j	] . !				·	HSCKAVTDEEPFLIFANRYYLRKLNLDGSNY
	[			Y		TLLKQGLNNAVALDFDYREQMIYWTDVTTQ GSMIRRMHLNGSNVQVLHRTGLSNPDGLAV
				E7E		DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL
				•		VSSGLREPRALVVDVQNGYLYWIDWGDHSL
						IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE
						RIYWADAREDYIEFASLDGSNRHVVLSQDIPH
1	[					IFALTLFEDYVYWTDWETKSINRAHKTTGTN
1	•					KTLLISTLHRPMDLHVFHALRQPDVPNHPCK
						VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD
						GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN
						PAFICDGDNDCQDNSDEANCDIHVCLPSQFK
i 1	1					CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE
						VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD
						GSDEPANCTQMTCGVDEFRCKDSGRCIPARW
						KCDGEDDCGDGSDEPKEECDERTCEPYQFRC
	ŀ					KNNRCVPGRWQCDYDNDCGDNSDEESCTPR
1						PCSESEFSCANGRCIAGRWKCDGDHDCADGS
				•	·	DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFOCNNT
		_				LCKPLAWKCDGEDDCGDNSDENPEECARFV
					·	CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD
		' I				GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL
						SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT
						NASICGDEARCVRTEKAAYCACRSGFHTVPG
]						QPGCQDINECLRFGTCSQLCNNTKGGHLCSC
						ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS
]						LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR
1 1		ļ [				ROIDRGVTHLNISGLKMPRGIAIDWVAGNVY
						WTDSGRDVIEVAQMKGENRKTLISGMIDEPH
] .						AIVVDPLRGTMYWSDWGNHPKIETAAMDGT
1						LRETLVQDNIQWPTGLAVDYHNERLYWADA
1					ļ	KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV
			*			FEDYTYGVTYINNRVFKIHKFGHSPLVNLTGG
						LSHASDVVLYHQHKQPEVTNPCDRKKCEWL
						CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP
1						PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC
		· •	[			QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGPTGPKCTQQVCAGYCANNSTCT
		1				VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE
		İ				NFGTCQMAADGSRQCRCTAYFEGSRCEVNK
<b>j</b>						CSRCLEGACVVNKQSGDVTCNCTDGRVAPS
		l J			l	CLTCVGHCSNGGSCTMNSKMMPECQCPPHM
						TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL
] .					. <b>i</b>	VAGVVFWYKRRVQGAKGFQHQRMTNGAM
L						NVEIGNPTYKMYEGGEPDDVGGLLDADFAL
·						

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D-Aspartic Acid, E-Ghitamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	scq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496 914	correspondi	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first		Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
	`			peptide	sequence	/=possible nucleotide deletion, \=possible
i	ĺ	l	ĺ	sequence	Ì	nucleotide insertion
	ļ			Sequence		DPDKPTNFTNPVYATLYMGGHGSRHSLASTD
						EKRELLGRGPEDEIGDPLA
374	1724	A	3187	191	1815	CLELASAGKIPEESKALSLLAPAPTMTSLMPG
	;					AGLLPIPTPNPLTTLGVSLSSLGAIPAAALDPNI
		ļ		1		ATLGEIPOPPLMGNVDPSKIDEIRRTVYVGNL
1		j	ļ		j	NSQTTTADQLLEFFKQVGEVKFVRMAGDET
1		1				QPTRFAFVEFADQNSVPRALAFNGVMFGDRP
			1		1	LKINHSNNAIVKPPEMTPQAAAKELEEVMKR
			İ		İ	VREAQSFISAAIEPGWLHSTSLCNDFLGCF*RR
[		1	ŀ			RMYRE*APCTICGTFHLCLIINWDL*LF*AYTA
	•		Ì			K*FFPPRVWKEQ*KKRR\RSRSHTRSKSRSSSK
			į			SHSRRKRSQSKHRSRSHNRSRSRQKDRRRSK
ł		ł				SPHKKRSKSRERRKSRSRSHSRDKRKDTREKI
ļ						KEKERVKEKDREKEREREKEREKERGKN
1						KDRDKEREKDREKDKEKDREREREKEHEKD
						RDKEKEKEQDKEKEREKDRSKEIDEKRKKDK
			'			KSRTPPRSYNASRRSRSSSRERRRRRSRSSSRS
						PRTSKTIKRKSSRSPSPRSRNKKDKKREKERD HISERRERERSTSMRKSSNDRDGKEKLEKNST
						S
375	1725	A	3192	415	101	AHSSHQTRAILQEFQWDIIRHPPL\SPNLALSG
1		••	J.,	71.5	101	F\FPNLKKSLRGTHFSSVKK\TTLTWLNSQDP
1						WF/FFYP*SPDLQIPSSFRNGLNDWYHHSOKC
1						PDLDGAYVKK
376	1726	A	3199	931	418	GV*WCDLGSPQPPPPGFKQFCLGRSSSWDYR
					1	HVPPHPANFVFLLETGFLHAGQAGL\GDPPAS
						ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI
						K
377	1727	A	3201	274	1285	KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM
						GYKLQDLTDVQIMARLQEESLRQDYASTSAS
						VSRHSSSVSLSSGKKGTCSDQEYDQYSLEDEE
1						EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC
[						RRSPSSQYFPSNNYQQQQYYSPQAQTPDQQP
i	1					NRTNGDK/PPKKYA*PSPDAKYNCH**QH\SSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGOL
1						
						QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS   NMPLSNGLQLYSNTGIPIPNKAAASGIMGRS
				]		ALPRPSLAINGSNLPRSKIAOPVRSFLOPPKPL
1						SSLSTLRDGNWRDGCY
378	1728	A	3202	112	1789	VPGVTESRPSVLRGDHLFALLSSETHQEDPIT
						YKGFVHKV\ELDRVKLSFSMSLLSRFVGWG*
						PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW
						PMLFP\VAPRDVPLLPSDVKLKLYDRSLESNP
					•	EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT
						LVEAIKQVVKHLPKAHILACAPSNSGADLLC
						QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN
						WDAKKGEYVFPAKKKLQEYRVLITTLITAGR
						LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG
						LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL
						TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ
					·	FITKLLRNYRSHPTILDIPNQLYYEGELQACA
[ ]						DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD
						EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK
) l					1	GKARLSPRSVGVISPYRKQVEKIRYCITKLDR
)						ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP
						ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD
379	1770		2205	422	120	SNLRVWDGIRKPACLTNTSCHS
3/7	1729	A	3206	432	130	PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK
						*LSTREAXDSXPGRQIAXXRQGGKVETTTAL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine O=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	}	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		,	914 .	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	1	İ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	Ì			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		İ		peptide		/=possible nucleotide deletion, \=possible
	<del></del>	<u> </u>		sequence		nucleotide insertion
	[	ĺ				XKQSNNKGTRASSYXEPDAXEQWKFPHKKL
380	1730	A	3207	187	507	QLPGXTHE
300	1730	Α	3207	107	1 207	GGTGHPHPARPPLSGVGGCQCSHSKPWTAGS
1	}				]	PEQRDHPAPHKQIEAGQGLPGPQAWGG*KGP AXLLPGPGGGPGPVASLEARAQASSGVTPNG
		İ				GGRTYPYPTFSSGE
381	1731	A	3225	1	840	GTRPGHLPAPSDGFCV/HL*SIPSWGSF*GESL/
1		1 **	1 3223	•	040	EMQLITSLGLQEFDIARNVLELIYAQTLVWIGI
	·					FFCPLLPFIQMIMLFIMFYSKNISLMMNFQPPS
	1					KAWRASQMMTFFIFLLFFPSFTGVLCTLAITI
ļ		l				WRLKPSADCGPFRGLPLFIHSIYSWIDTLSTRP
						GYLWVVWIYRNLIGSVHFFFILTLIVLIITYLY
			<b>!</b>	0.0		WQITEGRKIMIRLLHEQINEGKDKMFLIEKLI
	İ		Ì			KLQDMEKKANPSSLVLERREVEQQGFLHLGE
L		l			·	HDGSLDLRSRRSVQEGNPRA
382	1732	Α	3238	256	38	LLMIKVSSTCFSCHLHHHHHHHHHHRHHOGHNS
1						LFFSLKSSSNSSTLPVYLSYNIILVFSKCLVFDF
						LFSNACL
383	1733	A	3241	1542	343	KGAPSFVRLYQYPNFAGPHAALANKSFFKAD
1						KVTMLWNKKATAVLVIASTDVDKTGASYYG
1			!			EQTLHYIATNGESAVVQLPKNGPIYDVVWNS
Į.						SSTEFCAVYGFMPAKATIFNLKCDPVFDFGTG
1						PRNAAYYSPHGHILVLAGFGNLILQI*AD/IMK
1 .						VWNVKNYKLISKPVASDSTYFAWCPDGEHIL
						TATCAPRLRVNNGYKIWHYTGSILHKYDVPS
						NAELWQVSWQPFLDGIFPAKTITYQAVPSEVP
1						NEEPKVATAYRPPALRNKPITNSKLHEEEPPQ NMKPQSGNDKPLSKTALKNQRKHEAKKAAK
						QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP
[				ſ		EIDKKIKNLKKKLKAIEQLKEQAATGKOLEK
1						NQLEKIQKETALLQELEDLELGI
384	1734	A	3242	3	678	IRSPAARSPGLETPTCLLFVIAAIAAVFVDSAIP
		1	•	_		RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP
1						LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS
						ASLVRATVRAVSKRKLQPTRAALTLTPSAVN
<b>!</b>			[			KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL
1						BYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL
		1				GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES
-002	1000					FNI
385	1735	Α -	3243	3190	664	VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL
	<b>[</b>					KEEEILPEPGSETPTVASEALAELLHGALLRR
] ]		I	l	Ī	į	GPEMGYLPGPPLGPEGGEEETTTTITTTTVTT
		ł	- 1		ŀ	TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL
		1	1			GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL
ļ ļ		J	)	1	}	VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT
		l		1		NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP
			ļ	ĺ		PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG
		ļ	- 1	l	İ	EETLICLNGTRPSWNGETPSCMASCGGTIHNA TI GRIVSPERGGAVGDNI TCRWYIE A AEGRAI
	l	l	- 1	j	ļ	TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL
	İ	i		1		HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL
				l		SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG
•	i	- 1		l		ALATESCLPGYALEPPGPPNAIECVDPTEPHW
1	ľ		ľ	ł	l	NDTEPACKAMCGGELSEPAGVVLSPDWPOS
	ŀ	l	1			YSPGQDCVWGVHVQEEKRILLQVEILNVREG
	i	ł	1	ļ.		DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS
			1	1	1	GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR
	1	.	1	l		NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ
]				l	ļ	CEPGYELLGSDILTCQWDLSWSAAPPACQKI

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide seq-	1	in USSN	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence	j	055N 09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	Londo		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ł		]	l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	Ì	peptide		/-possible nucleotide deletion, /-possible
<u> </u>	<u> </u>	<b> </b>	<u> </u>	sequence		nucleotide insertion
Í		ĺ	1	[		MTCADPGEIANGHRTASDAGFPVGSHVQYRC
				ļ		LPGYSLEGAAMLTCYSRDTGTPKWSDRVPKC ALKYEPCLNPGVPENGYQTLYKHHYQAGESL
ł		1		ĺ	1	RFFCYEGFELIGEVTITCVPGHPSQWTSQPPLC
						KVTQTTDPSRQLEGGNLALAILLPLGLVIVLG
			Ì	ļ		SGVYIYYTKLQGKSLFGFSGSHSYSPITVESDF
386	1736		20.50		-	SNPLYEAGDTREYEVSI
300	1/30	A	3250	5725	3984	GTSTVTMATKKHFSIILNLLGMLLKKDNQDT
						RKLLMTWALEVAVVMKKSETYAPLFCLPSF HKFCKGLLADTLVEDVNICLQACSSLHALSSS
						LPDDLLQRCVDVCRVQLVHRGTCIRQAFGKL
						LKSIPLGVFLSNNNHTEIQEISLALRSHMSKAP
						SNTFHPQDFSD/VISFILYGNSHRTGKDNWLE
}						RLFYSCQRLDKRDQSTIPRNLLKTDAVLWQW
						AIWEAAQFTVLSKLRTPLGRAQDTFQTIEGIIR
						SLAGHTLNPDQDVSQWTTADNDEGHGNNQL RLVLLLQYLENLEKLMYNAYEGCANALTSPP
}						KVIRTFLYTNRQTCQDWLTRIRLSIMRVGLLA
1	İ					GQPAVTVRHGFDLLTEMKTTSLSQGNELEVSI
						MMVVEALCELHCPEAIQGIAVWSSSIVGKHL
						LWINSVAQQAEGRFEKASVEYQEHLCAMTG
						VDCCISSFDKSVLTLASAGCKSASLKHCLNGE
						SRKSVLSKPTDSSPEVINYLGNKACECYISTA DWAAVQEWQNAIHDLKKSTSSTSLNLKADF
						NYIKSLSSFESGKFVECTEQLELLPGENINLLA
						GGSKEKIDMKKLLRNM
387	1737	A	3255	380	76	MDIFLYNCKYQVQTEI*NSIQHIMA\SKKLSRF
		İ				LKYVHNL*AENYKTLMK*INEDLNKQRDVPY
) )						S*TARLNKMSIPTKTIFRFKAIYIKIPATYFIET NMO
388	1738	A	3260	685	428	PQWLGLQVYALPPANFVFFVEMRSTILAQTG
						PELLDSSDLPASASKSAGITCMSHHARTLSLK
ļ						*WPFCLSATQEKFC*PASEGVAW
389	1739	Α	3269	1	332	LDGYHTPIYMLNRIIRLPAAL*IISDQTGHALTI
]				j		LTRLETQMINADYQNKLTLDYLLTTDREVYE
1					•	PFNLTNYCLHIHNQRLGAYDLG*V*Q/KLAHV PVOV*HGFDPEAMFR
390	1740	A	3270	2	372	GRCHDQNKGKS\DGPDAQAEACGGESTYQEL
				1		LVNQNPIGQPLACRRLTRKIYEGIKKAVKPNH
	ļ		J	ļ		SPRGVKKVHKFVNKGEKGIMVLAGDTLGIGV
201	1741	<u> </u>	2000		100	YCLLPCMC*DRKLTYAHIPSTTDLGAGAGY
391	1741	A	3273	1	187	FFQEMLDIMKAISDMMGKCTYPVLKEDAPRQ
	1	- 1	ł			HVETFFQEELTRSQEGMKLGENFLMFAMPP DDSKESKGK*FFQEMLDIMKAISDMMGKCTY
] ]	ļ					PVLKEDAPRQHVETFFQVGINQKSRGHEVRR
						KFPDVCHAPR
392	1742	A	3281	901	521	FFFGDGVSPCRQAGV*WHDLDSLQNLPPGFK
l			1			RFSYLSLPSSW\DYRHVLPRQANFCIF/M*RRG
	)	ļ			J	FTMLARMVSIS*PRDLPALASQSAGITGVSHH
393	1743	A	3283	385	3	APPOMDFTFALLCFALKGCLPRQKEGGTLNLI
		1	2203	303	·	RNRSVVPEFVLLGLSAGPQTQTLLFVLFVVIC LLTVMGNLLLLVVINADSCLHTPMYFFLGQL
	1	ı	- 1			SFLDLCHSSVTAPKLLENLLSEKKTISVEGCM
	1	1	1			A*VFFVFATGGTESSLLAVMAYDRYVAIRTR
<u>                                     </u>						G
394	1744	A	3284	575	1054	CTKCKADCDTCFNKNFCTKCKSGFYLHLGKC
						LDNCPEGLEANNHTMECVSIVHCEVSEWNP
}	ŀ	}	ì		1	WSPCTKKGKTCGFKRGTETRVREIIQHPSAKG NLCPPTNETRKCTVQRKKCOKGERGKKGRE
						THE THE TAKE TO YOUR CONTROL TO THE TAKE THE TAK

SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nuci-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	l		ł	peptide		/=possible nucleotide deletion, \=possible
L				sequence		nucleotide insertion
				l		RKRKKPNKGESKEAIPDSKSLESSKEIPEQREN
205	1010	<u> </u>	2005	<u> </u>		KQQQ
395	1745	A	3286	1	340	RVLYVPSMGFCILVAHGWQKISTKSVFKKLS
٠ ,	ļ				ì	WICLSMVILTHSLKTFHRNWDWESEYTLFMS
						ALKVNKNNAKLWNNVGHALENEKNFERAL
396	1746	A	3293	1	172	KYFLQATHVQPDDIGAHMNVGR
370	1740	A	3293	1	172	GFRAVVMTVKTEAAKGTLTYSRMRGMVAIL
397	1747	A	3295	12	401	IAFMKQRRMGLNDFIQKIANNSYACKQ AEPACGASSCTPPSLRSSSSOSVGPLRPGRPL
351	1741	^	3293	12	401	WSEACAFL*AAAPQGPASPCCGLPSGFPRVW
	•	]				AQCCPPGGALRFPEGLGSVLSPRRCPOVSRGS
1	•	ļ	ľ			GLSAVPQEVPSGFLGPGLRACPQEAPSRFLRA
		<b>i</b> .		ļ	]	GLT
398	1748	A	3300	1912	2768	KQRRWQNIQRKGPKRYIVIAGNSQSHQPMIFS
						MLRKLPKVTCRDVLPEIRAICIEEIGCWMQSY
		1		]		STSFLTDSYLKYIGWTLHDKHREVRVKCVKA
						LKGLYGNRDLTARLELFTGRFKDWMVSMIV
				ł		DREYSVAVEAVRLLILILKNMEGVLMDVDCE
						SVYPIV*ASN*GLASAVGEFLYWKLFYPECEI
				1		RTMGGREQRQSPGAQRTFFQLLLSFFVESKSH
		l		j		SVTQAGVQWQFSAHRDLCLPGSSNSHVSASR
			( · (	ſ		VAGIAGAHRHTWLIYVFFSWRQGFAVLAGL
						VSNS
399	1749	A	3301	536	2391	LRSYGCKAPSRISHLHK\FLFLLLPSLLMGYSE
						SPPPITDSWAPFISLTHHVLSQSQSPLSSNCWI
						CLSTHTQ*FTALPADLLTWTQSNVSLHISYLAI
ĺ						PFLADSFLKPV/L*PGNSAKHLSFKLSSLSMVS
· [						GRAVALLHLIASGLTSIQTNTASSKPPIWGY\L
						STQTSFISPPPLCLSRTYPNPAHATMVGQVPQ SLCGLIFTL/RTPCRPSILHPNYKIISTSAWQKV
						LCFSGSPTIHTSLHLTTGSSFLSFHPIPGFPAAN
						SALYVSSLKGPPGKNVTIPSPVTGT*QPPHRGS
						N/RLTVDKDNFFLSPKPNSLHQLPSQ\TPYQAL
						TGAALAGSYPIWENENTLSWLPTFTYNFCLST
						PSLFFLCDTN*YLCLPANWSGTCTLVFQAPTI
						NILPPNQTILISVEASISSSPIRNKWALHLITLLT
						GLGITAALGTGIAGITTSITSYQTLFTTLSNTVE
					,	DMHTSITSLQRQLDFLVGVILQNWRVLDLLT
						TEKGGTCIYLQEECCFCVNESGIVHIAVRRLH
						DRAAEL*HQVADSWWQGSSLLRWIPWVAPF
<u> </u>			. 1			LGPLIFLFLLLMIGPCIFNLVSRFISQRLNCFIQ
						ASMQKHIDNIFHLCHV*YQSLRGNHSEAPEPR
	1000				4.50	P
400	1750	A	3303	2	453	THWRHSSGVPGSTTARRRRELEIATSDNQE
	,					YYNRLCQEVTNRERNDQKMLADLDDLNRTK
į						KYLEERLIELLRDKDALWQKSDALEFQQKLS
	•					AEERWLGDTEANHCLDCKREFSWMVRRHHC
401	1761		2204	ļ., ——	626	RICGRIFCYYCCNNYVLSKHGGKKERCC
401	1751	A	3304	1	626	MAPQHSSLDDKVPQQASTVCFEFQDILQHSQ
						CTEHKDSLWGPGARSQPFGAHNTRLSPDSCP
ļ					ļ	EKIVLRALKDSRAGMPEQDKDPGVQENPDD
ļ <b>!</b>						QRRVPQGTGDAPSAFRPLWDNGGLSPFVSRP
						GPLERDLHAQRSEVTYNQRSQSSWMSSFPKR
						NAFVSPYSSMGQAQP/GLPKTNPIGESCCWEG
402	1752		2205	1670	122	LSLSTQILG*QKPSKYIPSLCKR
402	1/32	A	3305	1678	172	MELPSGPGPERLFDSHRLPGDCFLLLVLLLYA
						PVGFCLLVLRLFLGIHVFLVSCALPDSVLRRF VVRTMCAVLGLVARQEDSGLRDHSVRVLISN
ł			}			HVTPFDHNIVNLLTTCSTVSESEAESATGRFP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide	1	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	1	USSN	location	corresponding	
seq-	uence	1	09/496	correspondi	to last amino	
uence	1	1	914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, O=Glutamine, R=Arginine, S=Serine.
	i	l	1 /14	amino acid	of peptide	
		ľ	1 .	residue of	sequence	T=Threonine, V=Valine, W=Tryptophan,
1	1	1	İ	peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	l		sequence		/=possible nucleotide deletion, \=possible
	<del> </del>	<del> </del>	<del> </del>	sequence		nucleotide insertion
I	1	ĺ	ĺ		ĺ	GAQLKAPLSPLAFRMEDTEALPLTPILYPTCQ
	1					FFFVFLNIFLLAFSSPGSQPLLNSPPSFVCWSR
1	ł	l				GFMEMNGRGELVESLKRFCASTRLPPTPLLLF
	j	1	]			PEEEATNGREGLLRFSSWPFSIQDVVQPLTLQ
					ĺ	VQRTLVSVTVSDASWVSELL\WSLFVPFTVY
ļ.	ļ					QVRWLRPVHRQLGEANEEFALRVQQLVAKE
Į.	i	ł	1 1		}	LG\QTGTRLTPA\DKAEHMKRQRHPR\LRPQS
ľ		ŀ				AQSSFPPSPWVLSS/SDVQTGQTLGFREFKESF
		ŀ	}			CPHVAIGVFIPERPWPKTGCCKTLTIHLILL+G
	1					GPVSFSCPE\DIHPRGT*VPTQQASGLPSFPSYG
-	1 .	!				PARGGVL*HPSAQQPLTFA\KSS\WARAGRAL
403	1753	A	3307	44	447	QERKQ\ALYEYARRFTERRAPGGLD
203	1,33	^	ا /الادر	**	447	DPSPSLLAVALGLRAGERTRSGPGSSPSGGIS
	1					GGASAGLASSPECACGRSHFTCAVSALGECT
İ	1 .					CIPAQWQCDGDNDCGDHSDEDGCILPTCSPL
1						DFHCDNGKCIRRSWVCDSDNDCEDDSDEQD
404	1754		3311	400		CPPRECEED
404	1754	A	3311	409	1	PRHGWGRRVLGRDRPRLQKVKKSVKAIYIPG
						QDHVQNEEIYARVLDKFGSNFLSRDNADLGT
						AFVKFSTLTK*LSALLKNLLQGLSRNVIFTLDS
1						LLKGDLKGVKGDLKKPFDKAWKDYETKFAK
405	1755		2200		150	IEKEKREREWR
403	1/23	A	3322	12	458	AAVPVENPWDDPRVRPRVRIFTWEDCIAGQA
- 0.0	ŀ					KVLCNDSYGVTIDWSPKGAFIRLTSQSVGNG
1						HPASKENDQMVDTIKNTTKVPIIWTYGDMVE
						PRPQMIRPAVGAKHKELWKILMALKKIK\IWE
406	1756	A	3324	1	426	GKYTKPSQYNPNYMLELAHNDSVW
- 400	1750	^	3324	1	420	LSMLSTISTEHRLSVLWPIWYCCHCPTHLSAV
1			i			MCVLLWALSLLQSILEWMFCSFLFSDVDSDN
!						WCQILDFLTAVWLIFLILVLCGFTLVLLVRIIC
1		ĺ				GSQKMPLTRLYVTILLTGLVFLFCSLPLSIQ*F
407	1757	A	3328	213	1841	LLYWIEKDLDDL
70,	1,3,		3326	213	1041	SGDLSPAELMMLTIGDVIKQLIEAHEQGKDID
						LNKVKTKTAAKYGLSAQPRLVDIIAAVPPQY
1 1		- 1			ł	RKVLMPKLKAKPIRTASGIAVVAVMCKPHRC
{		i	1	1		PHISFTGNICVYCPGGPDSDFEYSTQSYTGYEP
i l			ľ			TSMRAIRARYDPFLQTRHRIEQLKQLGHSVD
				.		KVEFIVMGGTFMALPEEYRDYFIRNLHDALS GHTSNNIYEAVKYSERSLTKCIGITIETRPDYC
1		ı	ŀ	ł		MKRHLSDMLTYGCTRLEIGVQSVYEDVARD
1			ŀ			TNRGHTVKAVCESFHLAKDSGFKVVAHMMP
						DLPNVGLERDIEQFTEFFENPAFRPDGLKLYP
		l	i	. [	1	TLVIRGTGLYELWKSGRYKSYSPSDLVELVA
[ [		1	1	ĺ		RILALVPPWTRVYRVQRDIPMPLVSSGVEHG
[		1	ļ		· ·	NLRELALARMKDLGIQCRDVRTREVGIQEIH
		j	İ	J	J	HKVRPYQVELVRRDYVANGGWETFLSYEDP
		- 1		1	1	
	l				1	DQDILIGLLRLRKCSEETFRFELGGGVSIVREL
				1		HVYGSVVPVSSRDPTKFQHQGFGMLMEEA
	' I	- 1		1		ERIAREEHGSGKIAVISGVGTRNYYRKIGYRL
408	1758	A	3335	3	467	QGPYMVKMLK
	.,,,,	^	رووو	<i>-</i>	707	AIASPRAAGIRHELTSTMAAGKNKRLTKGGK
1	j	ŀ			1	KGAKKKAV/DNIINIGKTLVTRTQRTKIASDG
	ł	- 1	- 1		1	LKGRVFEESLADLQND\TDGYLLRVI*VAFTT
J. l	ĺ		l			ERTNQI/REVFNKLIPDSIGKDIEKACQSIYPLH
409	1759	_	2220		1000	DDFARKVKMLKKPKFELRKLMELHGEGSS
""	1/39	A	3338	7	1252	PRWRNSARDEILLSFPQNYYIQWLNGSLIHGL
	J	ľ		Į.		WNLASLFSNLCLFVLMPFAFFFLESEGFAGLK
[				ł		KGIRARILETLGMLLLLALLILGIVWVASALID
<b>∟</b>						NDAASMESLYDLWEFYLPYLYSCISLMGCLL

SEQ ID NO: of nucl- ectide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning mucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLLCTPVGL\SRMFTVMGQLLVKPTILEDLDE QIYIITLEEEALQRPTKWAVFIRW/KYNIMELE QELENVKTLKTKLERRKKASAWERNLVYPA VMVLLLIETSISVLLVACNILCLLVDETAMPK
						GTRGPGIGNASLSTFGFVGAALEIILIFYLMVS SVVGFYSLRFFGNFTPKKDDTTMTKIIGNCVS ILVLSSALPVMSRTLGITRFDLLGDFGRFNWL GNFYIVLSYNLLFAIVTTLCLVRKFTSAVREE LFKALGLHKLHLPNTSRDSETAKPSVNGHQK AL
410	1760	A	3339	127	1433	GSHRFSLASPLDPEVGPYCDTPTMRTLFNLL WLALACSPVHTTLSKSDAKKAASKTLLEKSQ FSDKPVQDRGLVVTDLKAESVVLEHRSYCSA KARDRHFAGDVLGYVTPWNSHGYDVTKVFG SKFTQISPVWLQLKRRGREMFEVTGLHDVDQ GWMRAVRKHAKGLP*CLGSCLRTGLTMISG/ YVLDSEDEIEELSKTVVQVAKNQHFDGFVVE VWNQLLSQKRVGLIHMLTHLAEALHQARLL ALLVIPPATTPGTDQLGMFTHKEFEQLAPVLD GFSLMTYDYSTAHQPGPNAPLSWVRACVQV LDPKSKWRSKILLGLNFYGMDYATSKDAREP VVGARYIQTLKDHRPRMVWDSQVSEHFFEY KKSRSGRHVVFYPTLKSLQVRLELARELGVG VSIWELGQGLDYFYDLL*VGIAASAVDVFFSK
411	1761	<b>A</b>	3342	74	2701	VATRKLAKGFTQFAKMTEGTKKTSKKFKFK FKGFGSFSNLPRSFTLRRSSASISRQSHLEPDTF EATQDDMVTVPKSPPAYARSSDMYSHMGTM PRPSIKKAQNSQAARQAQEAGPKPNLVPGGV PDPPGLEAAKEVMVKATGPLEDTPAMEPNPS AVEVDPIRKPEVPTGDVEEERPPRDVHSERAA GEPEAGSDYVKFSKEKYILDSSPEKLHKELEE ELKLSSTDLRSHAWYHGRIPREVSETLVQRN GDFLIRDSLTSLGDYVLTCRWRNQALHPKIN KVVVKAGESYTHIQYLFEQESFDHVPALVRY HVGSRKAVSEQSGAIIYCPVNRTFPLRYLEAS YGLGQGSSKPASPVSPSGPKGSHMKRRSVTM TDGLTADKVTRSDGCPTSTSLPRPRDSIRSCA LSMDQIPDLHSPMSPISESPSSPAYSTVTRVHA APAAPSATALPASPVARRSSEPQLCPGSAPKT HGESDKGPHTSPSHTLGKASPSPSLSSYSDPDS GHYCQLQPPVRGSREWAATETSSQQARSYGE RLKELSENGAPEGDWGKTFTVPIVEVTSSFNP ATFQSLLIPRDNRPLEVGLLRKVKELLAEVDA RTLARHVTKVDCLVARILGVTKEMQTLMGV RWGMELLTLPHGRKLRLDLLERFHTMSIML AVDILGCTGSAEERAALLHKTIQLAAELRGT MGNMFSFAAVMGALDMAQISRLEQTWVTLR QRHTEGAILYEKKLKPFLKSLNEGKEGPPLSN TTFPHVLPLITLLECDSAPPEGPEPWGSTEHGV EVVLAHLEAARTVAHHGGLYHTNAEVKLQG FQARPELLEVFSTEFQMRLLWGSQGASSSQA RRYEKFDKVLTALSHKLEPAVRSSEL
412	1762	A	3347	1	898	IDRAAECRTKPLPMAVSIRGNADSIVACLVLM VLYLIKKRLVACAAVFYGFAVHMKIYPETYI LPITLHLLPDRDNDKSLRQFRYTFQACL*ELL KRLCNRTALMFVAVAGLTFFALSFGFYYEYG WEFLEHTYFYHLTRRDIRHNFSPYFYMLYLT AESKWSFSLGIAAFLPQLILLSAVSFAYYRDL VFCWFLHTSIFVTFNKVCTSQYFLWYLCLLPL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide	1 200	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ſ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	uence		09/496		to last amino	
seq-	uciice		1	correspondi		M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Ghutamine, R=Arginine, S=Serine,
<b>i</b> .	ĺ	ĺ	ľ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
	}					VMPLVRMPWKRAVVLLMLWFIGQAMWLAP
1	1	l	ł		ł	AYVLEFQGKNTFLFIWLAGLFFLLINCSILIQII
1	l		1			SHYKEEPLTERIKYD
413	1763	A	3361	3	474	PIPVRWNSLEGRLLRGYEQHANDGKDYISRN
į		1		•		*DLRSWTAADMAAQIIKRKWEAEEFAEQIKA
İ						YLEGTCVER/LRTHLENGKETLQLTEQSSQPTI
1		l	l		į	PIVGIVAGLVLLGAVVTGAVVSAVMCRKKNS
1	ŀ	i	<b>(</b>			GHFLPTDRVSYSEAASSDHAQGSDVSLTACK
<b>!</b>						V V
414	1764	A	3363	1488	453	HQILELKKKILKTYNPDYDEDLVQEASSEDVL
414	1704	^	3303	1400	433	CVID GOVERNOUS GROUPS AND AND AND AND AND AND AND AND AND AND
						GVHMVDKDTERDIEMKRQLRRLRELHLYST
]		J				WKKYQEAMKTSLGVPQRERDEGSLGKPLCP
						PEILSETLPGSVKKRVCFPSEDHLEEFIAEHLP
1	i					EASNQSLLTVAHADAGTQTNGDLEDLEEHGP
1		l	i I			GQTVSEEATEVHMMEGDPDTLAELLIRDVLQ
						ELSSYNGEEE/DPEEVKTSLGVPQRGDLEDLE
		ŀ		·		EHVPGQTVSEBATGVHMMQVDPATLAKSDL
1	ł					EDLEEHVPEQTVSEEATGVHMMQVDPATLA
						KQLEDSTITGSHQQMSASPSSAPAEEATEKTK
					. /	VEEEVKTRKPKKKTRKPSKKSRWNVLKCWD
1			1			IFNIF
415	1765	Α	3369	431	315	IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP
						PSFSPSL
416	1766	A	3373	42	651	RQEKMGLGEIGASGVLRSMLKERKKQNMKG
'''	-1.00		55,5		051	NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS
						APGLCSPLHPLQPQQEASTCPSGTLQGREKAA
[. [						PGQGRPLCSLWAGGAGA\PGERGAEGRGPSD
		ľ				QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL
]	•					RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG
417	1767	A	3382		2061	PKPQCSLPSQVPGDTH
417	1767	A	3382	2	2001	EAQDPRACGPDAGGRFAARDAPGNSLRPPPS
						SPP/GWPGQLRLLPRVPGSELRCGKPERGRLP
						ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT
					1	WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP
1						KNAGRCGGA\RGPRLSLGPPPGPPPAPALPAR
						ASAGAGAAAAALAVGGVRGAGGARGTGGY
						GHCSGR/PTGRTGPGPQGPGPPMPARPR*AS\S
				' l	ı	TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV
				ļ		HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP
				1		R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP
				1		SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY
				I		PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH
						SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP
					×	PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC
						SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD
					İ	HSSCEGHPDLHAGREMPAAPGLSELERVRFT
			[	j		VGCGGLASGISSASVSGLSPNRAGGPGQGDW
			· · · · · · · · · · · · · · · · · · ·	ĺ		
				•		EMYPVSWQTQESGGQG/SPKTGR*VGMLQA
,				.		GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS
						GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS
						Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA
						KGIGQLSEIPLLNVETAFWSMWVTYFRK
418	1768	Α	3398	304	2121	EEEEEEDEDDDDNNEEEEFECYPPGMKVQV
			1			RYGRGKNQKMYEASIKDSDVEGGEVLYLVH
			[	1		YCGWNVRYDEWIKADKIVRPADKNVPKIKH
				ļ		RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN
					l	PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI
			1	1	i	EITSILNGLQASESSAEDSEQEDERGAQDMDN
			į			NGKEESKIDHLTNNRNDLISKEEONSSSLLEE
						THE TAKE TO THE TAKE

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi	corresponding to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine,
	0			amino acid residue of peptide	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
		ļ <u>.</u>	<u> </u>	sequence		nucleotide insertion
						NKVHADLVISKPVSKSPERLRKDIEVLSEDTD YEEDEVTKKRKDVKKDTTDKSSKPQIKRGKR RYCNTEECLKTGSPGKKEEKAKNKESLCMEN
*						SSNSSSDEDEEETKAKMTPTKKYNGLEEKRK SLRTTGFYSGFSEVAEKRIKLLNNSDERLQNS
						RAKDRKDVWSSIQGQWPKKTLKELFSDSDTE AAASPPHPAPEEGVAEESLQTVAEEESCSPSV
						ELEKPPPVNVDSKPIEEKTVEVNDRKAEFPSS GSNFSA*IPLPYLHLNRLHQSL*QKGSRQQSS
						VTVSEPLAPNQEEVRSIKSETDSTIEVDSVAGE LQDLQSERE*LASRF*CQCELKQ**SARTRTS*
	·					KSLYRSEKSERCSGRRKFIKKAEKKP*SNSGK OOKEGK
419	1769	A	3399	206	463	QRECLSIHIGQAGIQIGDACWELYCLEHGIQP NGVVLDTQQDQLENAKMEHTNASFDTFFCE
420	1770	Α	3408	1010	685	TRAGKHVPRALFVDLEPTVIDGIR
	1,70		3400	1010	083	RRLSFFF*IWSSVLVTQARVQWRDLGSPQPLP PGFKRFSCLSLPSSWDYRHPSPRPVNF/HVFLV
						VMGFHHVGQAGLELLTSGDLPALASQSARIT GVNHCAQPRGHFH
421	1771	A	3409	355	1326	ADSNLIESCWQELGLGPWGGDWRVEQVGAS ASLRFPREVCSIRFLFTAVSLLSLFLSAFWLGL
						LYLVSPLENEPKEMLTLSEYHERVRSQGQQL QQLQAELDKLHKEVSTVRAANSERVAKLVF
		·				QRLNEDFVRKPDYALSSVGASIDLQKTSHDY ADRNTAYFWNRFSFWNYARPPTVILEPHVFP
						GNCWAFEGDQGQVVIQLPGRVQLSDITLQHP PPSVEHTGGANSAPRDFAVFFLLSFFTHOGLO
						VYDETEVSLGKFTFDVEKSEIQTFHLQNDPPA AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT
422	1772	A	3412	2	421	SEGAEGSAQGPH
722	1772	^	3412	2	421	EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN YLVSCSMRSPESGKGEPGTARDYTPMGRPPP
		. }				PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP
423	1773	A	3420	91	706	VDTAGAPASPGPDVCE DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/
	. '		1			RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS
		ŀ			Ì	KTHLPGFVEQAEALKAKGVQVVACLSVNDA
	1					FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA
424	1774	A	3421	4	7688	LNVEPDGTGLTCSLAPNIISQL RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ
	Ì	. :				FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV
	ľ		į			SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY
•	[.				ĺ	LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR
		f				DHGTPALTASASVSVTALDVNDNNPTFTOPE
	İ			l		YTVRLNEDAAVGTSVVTVSAVDRDAHSVTTY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE
					1	RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE
				Ì	Ì	DTGENARITYFMEDSIPQFRIDADTGAVTTQA ELDYEDQVSYTLAITARDNGIPQKSDTTYLEI
					.	LVNDVNDNAPQFLRDSYQGSVYEDVPPFTSV
L <u>.</u>						LQISATDRDSGLNGRVFYTFQGGDDGDGDFI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alaninc O=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Dadamatic Acid E-Obst.
nucl-	peptide					D=Aspartic Acid, E=Glutamic Acid,
eotide		i	in Licent	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				emino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	]	/=possible nucleotide deletion, \=possible
1	]			sequence		nucleotide insertion
		<del></del>		- Januar	<del></del>	
1						VESTSGIVRTLRRLDRENVAQYVLRAYAVDK
1 .						GMPPARTPMEVTVTVLDVNDNPPVFEQDEFD
1						VFVEENSPIGLAVARVTATDPDEGTNAQIMY
1						QIVEGNIPEVFQLDIFSGELTALVDLDYEDRPE
						YVLVIQATSAPLVSRATVHVRLLDRNDNPPV
1						LGNFEILFNNYVTNRSSSFPGGAIGRVPAHDP
1						DISDSLTYSFERGNELSLVLLNASTGELKLSR
1						ALDNNRPLEAIMSVLVSDGVHSVTAQCALRV
1						TUTDEMLTHSTTLRLEDMSPERFLSPLLGLFIQ
1					Ì	AVAATLATPPDHVVVFNVQRDTDAPGGHILN
						VSLSVGQPPGPGGGPPFLPSEDLQERLYLNRS
ŀ						LLTAISAQRVLPFDDNICLREPCENYMRCVSV
[						LRFDSSAPFIASSSVLFRPIHPVGGLRCRCPPGF
1					•	TGDYCETEVDLCYSRPCGPHGRCRSREGGYT
4						CLCRDGYTGEHCEVSARSGRCTPGVCKNGGT
1						CVNLLVGGFKCDCPSGDFEKPYCQVTTRSFP
1						AHSFITFRGLRQRFHFTLALSFATKERDGLLL
1						YNGRFNEKHDFVALEVIQEQVQLTFSAGEST
						TTVSPFVPGGVSDGQWHTVQLKYYNKPLLG
1	1					QTGLPQGPSEQKVAVVTVDGCDTGVALRFGS
						VLGNYSCAA\QGTQGGSKKSLDLTGPLLLGG
1			l			VPDLPESFPVRMRQFVGCMRNLQVDSRHIDM
1				ļ		ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC
1						VNQWDAFSCECPLGFGGKSCAQEMANPQHF
]				l		LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD
1		ď	1	ŀ	İ	GVLLQAITRGRSTITLQLREGHVMLSVEGTGL
		- 1				QASSLRLEPGRANDGDWHHAQLALGAIGGP
1		1		j		
1		ļ	1			GHAILSFDYGQQRAEGNLGPRLHGLHLSNITV
1	[	- 1	Í		l l	GGIPGPAGGVARGFRGCLQGVRVSDTPEGVN
		1				SLDPSHGESINVEQGCSLPDPCDSNPCPANSY
		l		j		CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC
1		j		J	l	EHQSVCTRKPSAPHGYTCECPPNYLGPYCET
						RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC
		ľ		l		NKTSGECHCKENHYRPPGSPTCLLCDCYPTG
		Į				SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF
1 1		}	ļ	j		AEVTTNGCEVNYDSCPRAIEAGIWWPRTRFG
		1		l		LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF
		l				NCTSITFSELKGFAERLQRNESGLDSGRSQQL
1 . 1			1	İ		ALLLRNATQHTAGYFGSDVKVAYQLATRLL
		l	ł	ļ		AHESTQRGFGLSATQDVHFTENLLRVGSALL
]		ĺ			•	DTANKRHWELIOOTEGGTAWLLOHYEAYAS
	[	ĺ	[	1		ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN
						FAGAKLPRYEALRGEQPPDLETTVILPESVFR
1 1		j				ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ
1 1			ļ	J		GEAVACUTVDTI AGI I DIDTERDENDE DE COMP
		ľ	l	l		GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK
] ]				ļ		RPIINTPVVSISVHDDEELLPRALDKPVTVQFR
j			ı	1		LLETEERTKPICVFWNHSILVSGTGGWSARGC
1 1		j	]	J	}	EVVFRNESHVSCQCNHMTSFAVLMDVSRRE
- 8	<b> </b>	l	·	l		NGEILPLKTLTYVALGVTLAALLLTFFFLTLL
]		l	l		1	RILRSNQHGIRRNLTAALGLAQLVFLLGINQA
		l	l		}	DLPFACTVIAILLHFLYLCTFSWALLEALHLY
} !		ŀ	ļ	. 1	i	RALTEVRDVNTGPMRFYYMLGWGVPAFITG
1 1		l		-	ŀ	LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP
		l	ł		1	VAFAVSMSVFLYILAARASCAAQRQGFEKKG
1 1		ļ	1		ļ	PVSGLQPSFAVLLLLSATWLLALLSVNSDILL
1 1		l	ł	ł	ŀ	ELIVI EVICNOLUCCHELLI GARAR GALLERANONITE
'		İ	j	1		FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK
'.		j		1	1	LACSRKPSPDPALTTKSTLTSSYNCPSPYADG
] i		l		į	.	RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR
		l			1	EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ
L		l				H\DS*TRDFDSDLSLEDDQSGSYASTHSSDSEE
					1	

	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide	l	in	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	seq-		USSN 09/496	correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine,
uence	ualce		914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
400,00		ĺ	/ / /	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			peptide		/=possible nucleotide deletion. \=possible
L		L		sequence		nucleotide insertion
						EEEEEERAAFPGEQGWDSLLGPGAERLPLHS
	i	ļ	1		0	TPKDGGPGPGKAPWPGDFGTTAKESSGNGAP
					1	EERLRENGDALSREGSLGPLPGSSAQPHKGIL
					Ì	KKKCLPTISEKSSLLRLPLEQCTGSSRGSSASE
l		]			].	GSRGGPPSRPPPRQSLQEQLNGVMPIAMSIKA
425	1775	A	3429	155	1417	GTVDEDSSGSEFLFFNFLH
725	1,,,,	<b> </b> ^	3729	133	1417	GEPAVQSCDCGCTQRSCPWLLVAPGLLSSSSS RAASVREAEDAPLQPASIHPVSQGSRGPEGSL
						GSAECLPGDPLGARRATRAHSPVPGPPPSLPA
						AGTAVKRGLQPG*GA/GATSTPGTGAATGGL
	1					CGPAWAAPSAVGPCCCCPSISTTPSQMRSARP
			ĺ			SLGCLPSWAS\PGTEHPPGPQGPGPS*DLCSV*
					100	KREFQRGPWAGMVILHRISAADPARAPGPDS
						NLQSALQQPATGCSEPAAVYSPPIGLWGA++P
						EYG*PQHSLPG*TAPADR*P\AGIKDRVYSNSI
		•				YELLENGQRAGTCVLEYATPLQTLFAMSQYS
						QAGFSREDRLEQAKLFCRTLEDILADAPESQN
						NCRLIAYQEPADDSSFSLSQEVLRHLRQEEKE EVTVGSLKTSAVPSTSTMSQEPELLISGMEKP
						LPLRTDFS
426	1776	Α	3431	1662	369	AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR
						SSQSLHCLSPRQPNHYLQALRAVGGICQDYD/
						SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP
						AFGFGARIPPNFEVG*MRGKEGDGGRVSQAE
		- 1		-		KAGPHCSRLALTG\SHDFAINFDPENPECEGK
1		Α.				RGDFHLPRLPADTLHTGAQTPLPRAQLPVPST
1 1						HPRPVFT/EISGVIASYRRCLPQIQLYGPTNVAP IDNRVAEPAQREQSTGQATKYSVLLVLTDGV
			i			VSDMAETRTAIVRASRLPMSIIIVGVGNADFS
·	1		- 1			
						DMKLLDGDDGPLKCPKGVPAARDIVOFVPFR I
1						DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD
			:			DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP+SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ
400	Letter					DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP
427	1777	A	3446 .	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP+SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW
427	1777	A	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP+SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA
427	1777	A	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP+SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL
427	1777	A	3446 .	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQPPPP
427	1777	A	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP+SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL
427	1777	A	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQPPPP PPPPPPQLPQPPPQAQPLLPQPPPPPPPPPP
427	1777	A	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQPPPP PPPPPPQLPQPPPQAQPLLPQPPPPPPPPPP
427	1777	A	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRGGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQPPPP PPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK
427	1777	A	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRGGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP
427	1777	A	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRGGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAPPLPPPPPPPPP GPAVAEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPIT
427	1777	A	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRGGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSPQQQQQQQQQQQQQQQQQQQPPPP PPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPPP
427	1777	A	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV
427	1777	A	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFILCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL
427	1777	A	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFILCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL
427	1777	A	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI
427	1777	<b>A</b>	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFILCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL
427	1777	<b>A</b>	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS
427	1777	<b>A</b>	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFILCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWFFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG
427	1777	<b>A</b>	3446	79	9748	DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPQLPQPPPPQAQPLLPQPQPPPPPPPPPP
427	1777	<b>A</b>	3446	79		DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFILCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASEA
427	1777	<b>A</b>	3446	79		DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFILCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF
427	1777	<b>A</b>	3446	79		DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFILCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS
427	1777	<b>A</b>	3446	79		DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTLAAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFILCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	to last amino acid residue	l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
		<u>.</u>		amino acid residue of peptide	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
<u> </u>	<del>                                     </del>	<u> </u>		sequence		nucleotide insertion TTEYPEEQYVSDILNYIDHGDPQVRGATAILC
1		•				GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL
						ADCIPLLRKTLKDESSVTCKLACTAVRNCVM
						SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEL LETLAEIDFRLVSFLEAKAENLHRGAHHYTGL
						LKLQERVLNNVVIHLLGDEDPRVRHVAAASL
1						IRLVPKLFYKCDQGQADPVVAVARDQSSVYL
						KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE
						ALCLLSTAFPVCIWSLGWHCGVPPLSASDESR
].	·					KSCTVGMATMILTLLSSAWFPLDLSAHQDAL
						ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC
ļ		1				AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK
•						GKEKEPGEQASVPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA
ļ						NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL
1						ATLQDIGKCVEEILGYLKSCFSREPMMATVC
1						VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA
						SLRNMVQAEQENDTSGWFDVLQKVSTQLKT
1	1					NLTSVTKNRADKNAIHNHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL
						DSDQVFIGFVLKQFEYIEVGQFRESEAUPNIFF
						FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR
					Ì	KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLOO
1			1		ľ	CHKENEDKWKRLSRQIADIILPMLAKQQMHI
	i l					DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVOLVISSGILAH DVI LEOSTED
						VTPNIMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE
j						EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT
		ĺ				KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR
						ARSMITTHPALVLLWCQILLLVNHTDYRWW
						AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA
				ļ		AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVODFISAVHRNS
		.				AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI
						HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILAANLQSSMAQLPMEELNRIOEY
			ļ			LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS
1		l	j		j	PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK
	. !		ŀ		}	SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGOKSALFEA
1		I	į		1	AREVILARVSGTVQQLPAVHHVFQPELPAEP
		ł		ł	ł	AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL
					1	SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL
				İ		WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG
				}		EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA
1		ľ	ł	ł	1	FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG
	•					WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR
				ļ	,	INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT
					.	VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK
			1		İ	LSIRGIVEQEIQAMVSKRENIATHHLYQAWD
					ļ	PVPSLSPATTGALISHEKLLLQINPERELGSMS YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE

						AARC*PCT/SRPT*STLSPTQAAWWCRPSRRQ QRGEASTGGASGRRCGSCFQV
731	1/01	^	34/4	1		FRPAPGHVQP*GGSSAAAGGGLLSHPRPCQQ PCPPAPAPSRPRSLGSLGQRVPAALATAAQEL PATLGGDGGKPALTAGEAALPGLHRSGVPAA
431	1781	A	3474	1	441	MSLENDKILQIITELIKTENN FRPAPGHVQP*GGSSAAAGGGLLSHPRPCQQ
	.		ļ		.	LREITLITPFPVLLFGGDIEVQHRERLLSIDGW IYFQAPVKIAVIFKQLRVLIDSVLRKKLENPK
		.	j			GLYDNVGKIIYTKSVDVTEKLACIVETAQGK AQVHPSSVNRDLQTHGWLLYQEKIRYARVY
	1	}				TYCRRNFLNRTSLLTLEDVKQELIKLVKAAGF SSSTTSTSWEGNRASQTLSFQEIALLKAVLVA
						VATLAAVMTEKSPFTTPIGRKDEADLAKSAL AMADSDHLTTYNAYLGWKKARQEGGYRSEI
.		. [			1	LTPLGQHLAALPVNVKIGKMLIFGAIFGCLDP
					Ì	PMDYSVPEILRVPLEELCLHIMKCNLGSPEDF LSKALDPPQLQVISNAMNLLRKIGACELNEPK
						PDVVFVIDTGRTKENKYHESSQMSSLVETFVS KASALQRQGRAGRVRDGFCFRMYTRERFEG
		ł			·	PGLAHIQQLYDLLSNDRRFYSERYKVIALHSI LSTQDQAAAFTLPPPGVRKIVLATNIAETGITI
						PHKINLDLILELLAYLDKSPOFRNIEGAVLIFL
	.		-			KDSEYCQKFLEEEEEVTINVTSKAGGIKKYQE YIPVQTGAHADLNPFYQKYSSRTQHAILYMN
:						DFLLIILKEILQKRSDLHLILMSATVDSEKFST YFTHCPILRISGRSYPVEVFHLEDIIEETGFVLE
						PGGRNSLCGYQIRMESRACESTRLLYCTTGV LLRKLQEDGLLSNVS/HMFIVDEV\HER\SVQS
						ASKCNIVCTQPRRISAVSLANRVCDELGCENG
·						HLSIQKLLKER\Q\QLPVFKHRDSIVETLKRHR VVVVAGET\GSGKSTQVPHFLLEDLLLNEWE
430	1780	A	3473	2802	270	SGETDSE FRMRIFLHCPWNQQMWKIWNLLETSLESCKA
						NAGRR*KGGGKIGTKQGAVRARKECRGEMA
						CPGSCDPQVLSASWM*VEHRSKFRPPF*NSTI PPES/RS*QGGTVQTGQHSSGREAGSWRARGR
						RFSGVLEPPLPSLKDGGRFPAWT*RSCSKSLR AAFTSQFFPSRRSRASPGSAP\GNGQNLTEQHP
429	1779	A	3464	583	3	DALDRRYLERCHPAAGGWVGEGE*ALCOKT/
					•	WSPCSASCGGGHYQRTRSCTSPAPSPGEDICL GLHTEEALCATQACPEGWS
						WGPWSSCSRDCELGFRVRKRTCTNPEPRNGG LPCVGDAAEYQDCNPQACPVRGAWSCWTS
428	1778	A	3449	3	430	NSRPSPSAALVEVLLRSGSTFPHTVSGGWAA
				·		VNLFCLVATDFYRHQIEEELDRRAFQSVLEV VAAPGSPYHRLLTCLRNVHKVTTC
	. ;					TGQSSMVRDWVMLSLSNFTQRAPVAMATWS LSCFFVSASTSPWVAAILPHVISRMGKLEQVD
						NKVIGEFLSNQQPYPQFMATVVYKVFQTLHS
						KEKVSPGRTSDPNPAAPDSESVIVAMERVSVL FDRIRKGFPCEARVVARILPQFLDDFFPPQDIM
		,				KLSVDRVNVHSPHRAMAALGLMLTCMYTG
						AFYLIENYPLDVGPEFSASIIQMCGVMLSGSE ESTPSIIYHCALRGLERLLLSEQLSRLDAESLV
						RSSHLPSRVGALHGILYVLECDLLDDTAKQLI PVISDYLLSNLKGIAHCVNIHSQQHVLVMCAT
						YLVPATCKAAAVLGMDKAVAEPVSRLLESTL
						LELYSRWILPSSSARRTPAILISEVVRSLLVVS DLFTERNQFELMYVTLTELRRVHPSEDEILAQ
						ADAPAPSSPPTSPVNSRKHRAGVDIHSCSOFL
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
uence	i	ļ	914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline.
nucl- eotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
NO: of	NO: of,	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.

SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
ectide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	j	914	ng to first	acid residue	Q=Ghutamine, R=Arginine, S=Serine,
ucaico		i	714	amino acid	of peptide	
i	[		ļ ·	residue of		T=Threonine, V=Valine, W=Tryptophan,
1	·		ŀ	peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1				possible nucleotide deletion, \=possible
432	1782	A	3478	sequence	22	nucleotide insertion
432	1/02	ΙΔ.	34/6	416	23	QLRRLTLPNFKTY/YSS*IIEIAWH**KNMQID
l	I	1				QWFRRESPEIDLCKYS*LSFDKEAKAIK/WKE
l .	1	)	ļ	}	ļ	CSLFNKWC/YKNWM/LHVQKKRI*VQTLHPS
		ĺ	ļ			QKLK\SKWIKDLNVECRITKLLDQEYPGDLGY
			<u> </u>			SRALNSGSR
433	1783	A	3504	1876	552	CLAPCSPQPEKNGMQPLLLLLPPLLYQQLLHS
1	i .					SLGAPGESTLLVRTSKLLVGLGLQLLVWLLL
1	l	l	1			QTRSLLALQLHLTSSAPLLAAPTAVCSCSRCS
· ·	ļ					APRSRCVARPAARTGLPTPAPASSPAPAASPA
						PAASPAPAESTA\PQPLILLPKP/PPAPGAPPPRP
						GAPPPRPAASPSPAASPAPPAASPVLTASPPLP
		1				AASPSPAASPAPPAASPVLTASPPLPAASPSPA
1	ł	1				ASPAPPAASPVLTASPPLPAASPALAASPVHT
•		i			•	ASPPVHVASPPVHTASPPVHVASPPVHTASPP
				}		VHVASPPVHTASPHVHVASPPVHVASPPVHV
}	]		]	·		ASPPVHTASPPVHVASPPVHTASPHVHVASPP
	l					VHTASPPVHVASPPVHVAYPPVHV
	l					ASPPVHVASPPVHVASPPVSCSGDSTSDCFPP
1	1					QPGAVFPHSLAPSLGGWSHLVAALP
434	1784	A	3516	142	590	GGVNRPRSETEQVKTPVLISSWDYRHPPPRPA
1	1,01	' '	3310	172	370	SFFVFLV*TGF\TALARMVLISWPCDLPTSASO
1 1						SAGITGVRHHA\RLLYFEQESHSVTQAGW\VQ
						MADE COLOR OF EDDI BECKE COCKE COCKE
						WHNLGSLQPLSLEDRLSPGVLGCSALCRSGV
435	1785	A	3529	1	3161	RTKFGINMVTSRERGTTRLPKEG
1 733	1765	^	3329	1	3101	MSLVRAALEALDELDLFGVKGGPQSVIHVLA
1						DEVQHCQSILNSLLPRASTSKEVDASLLSVVS
						FPAFAVEDSQLVELTKQEIITKLQGRYGCCRF
}						LRDGYKTPKEDPNRLYY/ENPAELKLFENIEC
						EWPLFWTYFILDGVFSGNAEQVQEYKEALEA
1 1						VLIKGKNGVPLLPELYSVPPDRVDEEYQNPHT
]						VDRVPMGKLPHMWGQSLYILGSLMAEGFLA
						PGEIDPLNRRFSTVPKPDVVVQVYPSLPHGCS
1 1				i		SKSPSHQCTIISIRTTRKITAPVSILAETEEIKTIL
1 1						KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF
						LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK
				!	l	LYDIRKTIFTFTPQFIDQQQFYLALDNKMIVE
}				j		MLRTDLSYLCSRWRMTGQPTTTFPISHSMLDE
] [		1	- 1	ſ		DGTSLNSSILAALRKMQDGYFGGARVQTGKL
		- 1	ļ			SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN
'						YDYLESGNWMNDYDSTSHARCGDEVARYL
			·	1	•	DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT
] ]		J	j			TCDLMSLVTKAKELHVQNVHMYLPTKLFOA
	İ	1	ĺ			SRPSFNLLDSPHPRQENOVPSVRVEIHLPRDO
		Ì		j		SGEVDFKALVLQLKETSSLQEQADILYMLYT
			-{			MKGPDWNTELYNERSATVRELLTELYGKVG
		-	l			EIRHWGLIRYISGILRKKVEALDEACTDLLSH
			j		ļ	QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA
	İ	ŀ		1	ľ	SEGDMSISILTQEIMVYLAMYMRTQPGLFAE
1			1	1		MFRLRIGLIIQVMATELAHSLRCSAEEATEGL
[	į	J	ł	ı		MNLSPSAMKNLLHHILSGKEFGVERSVRPTD
		. [	[	{	ſ	SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK
		}		i		QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW
	ļ	- 1	1	İ	]	ODDBDI DCVI MD/DD/CDAVA/MINA A CASA-
		ŀ	1		l	QRRRRLDGALNRVPVGFYQKVWKVLQKCH
	j	· 1	j	J	ì	GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL
	ļ	l	ĺ		. [	NRVPQPEYRQLLVEAIL\VLTMLADIE\HSIGS
		l	- 1			IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD
		l	1	ľ		PASGICTLLYDSAPSGRFGTMTYLSKAAATY
<del> </del>	1.562	I				VQEFLPHSICAMQ
436	1786	A	3546	73	202	COST TWELL DWW A DWI ODG! DCDWggggggg
		1	3370	13	393	CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL
<u> </u>		[	3340	/3	393	EQPDSCRPYGRSFYALEEKHVIFSLDVGETDN

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	поа	in NO:	nucleotide	nucleotide location	D=Aspartic Acid, E=Ghttamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	1	ŀ	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1		i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Joquasoo	/=possible nucleotide deletion, \=possible
ľ	ĺ	ľ	Î i	sequence		nucleotide insertion
_	<u> </u>	<del></del>	<u> </u>	304		KGKGKTIRGI*TFKGRKGGTYQREHDANPLA
	] ,					PXSARSCWMRKG
437	1787	A	3554	5157	2939	AVRAEPGLEELSSGLRAHSPSATTVCEPEAQG
	1,4,		555 :	,	-,,,	SASGCRYAAHPHWGLGGAAAAGGSWEPQPP
		[		,		RPVCEPAGRGKPHPPAAPRSPLLPGSRRRPHA
}	1	1			ļ	AQPGARARTSPPPASARNMAARPAATLAWSL
1	ł	·			ŀ	LLLSSALLREGCRARFVAERDSEDDGEEPVVF
1	i				ļ	PESPLQSPTVLVAVLARNAAHILPHFLGCLER
		ł			1	LDYPKSRMAIWAATDHNVDNTTEIFREWLK
						NVQRLYHYVEWRPMDEPESYPDE1GPKHWP
	0.0					TSRFAHVMKLRQAALRTAREKWSDYILFIDV
						DNFLTNPQTLNLLIAENKTIVAPMLESRGLYS
					ł	NFWCGITPKGFYKRTPDY\VQIREWKRTGCFP
				,		VPMVHSTFLIDLRKEASDKLTFYPPHQDYTW
			ļ '			TFDDIIVFAFSSRQAGIQMYLCNREHYGYLPIP
	[				1	LKPHQTLQEDIENLIHVQIEAMIDRPPMEPSQ
1			1			YVSVVPKYPDKMGFDEIFMINLKRRKGQGGD
]			1			RWLRTLYEQEIEVKIVEAVDGKALNTSQLKA
						LNIEMLPGYRDPYSSRPLTRGEIGCFLSHYSV
1						WKEVIDRELEKTLVIEDDVRFEHQFKKKLMK
						LMDNIDQAQLDWELIYIGRKRMQVKEPEKA
1.						VPNVANLVEADYSYWTLGYVISLEGAQKLV
	İ					GANPFGKMLPVDEFLPVMYNKHPVAEYKEY
						YESRDLKAFSAEPLLIYPTHYTGQPGYLSDTE TSTIWDNETVATDWDRTHAWKSRKQSRIYSN
1						AKNTEALPPPTSLDTVPSRDEL
438	1788	A	3563	130	527	IFFNSSSLFCRVFCLFLRWSFTLVAQARVQ*C
	.,,00		3203	150	52.	NLSSLQPLPPGFK*FSCLSPPRS*DYRRPPPRPA
						NFLYF**RQGFTVLGQAGLELLT/S/GDPPTSA
Į i						SQSAGITGVSHRAWPVHAISTHISLVKTRPSLT
1						TLG
439	1789	Α	3565	446	1834	LLQPAMRKSPGLSDCLWAWILLLSTLTGRSY
						GQPSLQDELKDNTTVFTRILDRLLDGYDNRL
						RPGLGERVTEVKTDIFVTSFGPVSDHDMEYTI
1						DVFFRQSWKDERLKFKGPMTVLRLNNLMAS
						KIWTPDTFFHNGKKSVAHNMTMPNKLLRITE
						DGTLLYTMRLTVR\AECPMAFGRDFPM\D\AH
				·		ACPLKFGSYAYTRAEVVYEWTREPARSVVV
1						AEDGSRLNQYDLLGQTVDSGIVQSSTGEYVV
						MTTHFHLKRKIGYFVIQTYLPCIMTVILSQVSF
						WLNRESVPARTVFGVTTVLTMTTLSISARNSL
			[			PKVAYATAMDWFIAVCYAFVFSALIEFATVN
				·		YFTKRGYAWDGKSVVPEKPKKVKDPLIKKN
					·	NTYAPTATSYTPNLARGDPGLATIAKSATIEP
	.					KEVKPETKPPEPKKTFNSVSKIDRLSRIAFPLL
440	1790	Ā	3568	1	360	FGIFNLVYWATYLNREPQLKAPTPHQ
440	טיפי/ נ	Α	2200	•	350	STSSCFPAAAAAIMREIVHLQAGQCGNQIGAK FWEVISDEHGIDPTGTYHGDSDLQLERINVYY
				İ		PWEVISDEHGIDPTGTYHGDSDLQLERINVYY   NEATGEAPVPSPTALRGPRGPCLG*RPPVPAG
441	1791	A	3569	2	1751	GKYVPRAVLVDMEPGTMDSV FVAVAGAVSGEPLVHWCTQQLRKTFGLDVS
TT	1/71	^	2202	4	1/21	EEIIQYVLSIESAEEIREYVTDLLQGNEGKKGQ
				i		FIEELITKWQKNDQELISDPLQQCFKKDEILDG
]			]			QKSGDHLKRGRKKGRNRQEVPAFTEPDTTAE
		1	ļ	1		VKTPFDLAKAQENSNSVKKKTKFVNLYTREG
			1			QDRLAVLLPGRHPCDCLGQKHKLINNCLICG
}		l	ì	,		RIVCEQEGSGPCLFCGTLVCTHEEQDILRGDS
						NKSQKLLKKLMSGVENSGKVDISTKDLLPH
-						QELRIKSGLEKAIKHKDKLLEFDRTSIRRTQVI
	<del>لــــــــــ</del> ــــــــــــــــــــــــــ	1				

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A-Alanine C-Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq~	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		[	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
	ŀ	ŀ		peptide	Sequence	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						DDESDYFASDSNQWLSKLERETLQKREEELR
						ELRHASRLSKKVTIDFAGRKILEEENSLAEYH
İ						SRLDETIQAIANGTLNQPLTKLDRSSEEPLGVL
	<b>]</b>					VNPNMYQSPPQWVDHTGAASQKKAFRSSGF
			,			GLEFNSFQHQLRIQDQEFQEGFDGGWCLSVH QPWASLLVRGIKRVEGRSWYTPHRGRLWIAA
						TAKKPSPQEVSELQATYRLLRGKDVEFPNDY
Ì						PSGCLLGCVDLIDCLSQKQFKEQFPDISQESDS
						PFVFICKNPQEMVVKFPIKGNPKIWKLDSKIH
						QGAKKGLMKQNKAV
442	1792	A	3576	1	2019	MPRSHTGERLCEGKEGSQCAENFSPNLSVTK
Į .	ļ.				]	KTAGVKPYECTICGKAFMRLSSLTRHMRSHT AIRAI\EKPYKCKEC\GRAFSLSQILSK\HERSH
		`		·		TGEKPYKCKQCGKTFIYHQPFQRHERTHIGEK
						PYECKQCGKALSCSSSLRVHERIHTGEKPYEC
						KQCGKAFSCSSSIRVHERTHTGEKPYACK\EC
				·		GKAFIS\TTSVLTHMITHNGDRPYKCKECGKA
						FIFPSFLRVHERIHTGEKPYKCKQCGKAFRWS
						TSIQIHERIHTGEKPYKCKECGKSFSARPAFRV HVRVHTGEKPYKCKECGKAFSRISYFRIHERT
						HTGEKPYECKKCGKTFNYPLDLKIHKRNHTG
						EKPYECKECAKTFISLENFRRHMITHTGDGPY
			,			KCRDCGKVFIFPSALRTHERTHTGEKPYECKQ
		•				CGKAFSCSSYIRIHKRTHTGEK\PYECKECGK
					. :	AFTYPTSFQGHMRMHTGEKPYKCKECGKAFS LHSSFR\RHTRIHNYEKPLEC*Q\CGKAFSVSTS
						LKKPMRNAQSDRKLY/KCEK*EKVFNSNRCF
						QSCENSH*REKSCQCK*YRKRDTR*FMYSQV
		•				PHNHVSVSNGPYR/CGSPIRLYNT*NISINRNL
						VAVVTP*CSTLFKCLWCWCKRAALSVV*/IVQ
						DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL   ANTVKPHLY
443	1793	A	3578	287	114	DFYERKFEOFIEGHKOIVNKWRDLLCSWKRK
	•					LSIIKKSVLQNNL*FSAASMRFQKVFF
444	1794	A	3582	3335	1909	HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT
						MTKVTLENFYSNLIAQHEEREMRQKKLEKV
						MEEEGLKDEEKRLRRSAHARKETEFLRLKRT RLGLEDFESLKVIGRGAFGEVRLVOKKDTGH
		i				VYAMKILRKADMLEKEQVGHIRAERDILVEA
						DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM
						MTLLMKKDTLTEEETQFYIAETVLAIDSIHQL
				٠ ا	.	GFIHRDIKPONLLLDSKGHVKLSDFGLCTGLK
						KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMOTGYN
						KLCDWWSLGVIMYEMLIGYPPFCSETPOETY
'					, [	KKVMNWKETLTFPPEVPISEKAKDLILRFCCE
		- 1				WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA
		1				AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE
<b> </b> .						TDYKNKDWVFINYTYKRFEGLTARGAIPSYM
445	1795	A	3584	1	6169	KAAK RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD
773	1,,,,	^	JJ04	1	0107	GGSRGKGEHFPYEQEIKFFAKVVLPLIDOYFK
		- 1		•		NHRLYFLSAASRPLCSGGHASNKEKEMVTSL
		- 1	}			FCKLGVLVRHRISLFGNDATSIVNCLHILGQT
			Ì			LDARTVMKTGLESVKSALRAFLDNAAEDLE
				l		KTMENLKQGQFTHTRNQPKGVTQIINYTTVA
				ľ		LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI
						LTSLYALGTSKSIYVERQRSALGECLAAFAGA FPVAFLETHLDKHNIYSIYNTKSSRERAALSLP
]		1				TNVEDVCPNIPSLEKLMEEIVELAESGIRYTO
	L					

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide insertion  MPHVMEVILPMLCSYMSRWWEHGPENNPER AEMCCTALNSEHMNTLLGNILKHYNNLGIDE GAWMKRLAVFSQPIINKVKPQLLKTHFLPLM EKLKKKAATVVSEEDHLKAEARGDMSEAEL LILDEFTTLARDLYAFYPLLIRFGDYNRAKWL KEPNPEAEELFRMVAEVFIYWSKSHNFKREE QNFVVQNEINNMSFLITDTKSKMSKAAVSDQ ERKKMKRKGDRYSMQTSLIVAALKRLLPIGL NICAPGDQELLALAKNRFSLKDTEDEVRDIIRS NIHLQGKLEDPAIRWQMALYKDLPNRTDDTS DPEKTVERVLDIANVLFHLEQKSKRVGRRHY CLVEHPQRSKKAVWHKLLSKQRKRAVVACP RMAPLYNLPRRAVNLFLQGYEKSWIETEEH YFEDKLIEDLAKPGAEPPEEDEGTKRVDPLHQ LILLFSRTALTEKCKLEEDFLYMAYADIMAKS CHDEEDDDGEEEVKSFEEKEMEKQKLLYQQ ARLHDRGAAEMVLQTISASK GETGPMVAAT LKLGIAILNGGNSTVQKMLDYLKEKKDVGF FQSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVQ ESISDFYWYYSGKDUDEQGQRNFSKAIQVA KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV VGFLHYFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFDMFILKLKDLTSSDTFKEYD PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENSTLDYEEFVKRFHEPAKDIGFNVA VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRGFIFDVVNEGGEKEKMELFVNFCED TIFFEMQLAAQISESDLNERSANKEESEKERPEE QGFRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKQMKVKKNTYKDMVTAFFSSYWSI FMTLHFYASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEGERKP LEAALPSEDLTDLEFFLAKMDLYFREPENSSENA KVTSLDSSSHRILAYHYVLEESSGYMEPTVRIL HTTLIFF SULTSSTFREFU QGFRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKQMKVKKMTVKDMVTAFFSSYWSI FMTLHFYASSPERFERIEGGEGEKKALFVNFCED TIFFEMQLAAQISESDLNERSANKEESEKERPEE QGFRMAFFSILTYRSALFALRYNILTLMRMLS LKSLKQMKVKKMTVKDMYTAFFSSYWSI FMTLHFYASSPERFILTGSLTGLIGGSUVEDAETKRIPHNYOUR FFREFERIEGGEGEKEEKEKEEFEKEEGEGGEKEEKAKEPVARK LEFDGLYTIPGPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKWMDKYGEFFGRDRISSLLG MDKAALDFSDAREKKRPKVSSSLSAVLNSID VKYQMWKLGVFTDNSFLYLAWYMTM
						FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN
446	1796	A	3592	1		AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/MIL PRLVSNSWTQAILLPRPPKMLGLQV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible mucleotide deletion, \=possible nucleotide insertion
447	1797	A	3598	1202	1070	LFVGGGPICPEGASGFAPGPAPAPRVGVDAEV GR*V*GAAASQGA/GSLRPRPTGPGHPGAWL QVWGAAAVCAGPAM*/AVRAKRGPRAG*EP NSPWRSGVLAA\RAVGAGPWP*P*PGCS*ARG PSSRSAPGLASGPAAPLLQGVHSSAGPLLCYI NGTLALGLKP**AWGWGEWRPKG
448	1798	A	3604	3115	557	FRRKGGGGPKDFGAGLKYNSRHEKVNGLEE GVEFLPVNNVKKVEKHGPGRWVVLAAVLIG LLLVLLGIGFLVWHLQYRDVRVQKVFNGYM RITNENFVDAYENSNS TEFVSLASKVKDALKL LYSGVPFLGPYHKESAVTAFSEGSVIAYYWSE FSIPQHLVEEAERVMAEERVVMLPPRARSLKS FVVTSVVAFPTDSKTVQRTQDNSCSFGLHAR GVELMRPTTPGFPDSPYPAHARCQWALRGD ADSVLSLTFRSFDLASCDERGRHLVTVYNTVL SPMEPHALVQLCGTYPPSYNLTFHSUSQNVL LITLITINTERRHPGFEATFFQLPRMSSCGGRL RKAQGTFNSPYYPGHYPPNIDCTWNIEVPNN QHVKVRFKFFYLLEPGVPAGTCPKDYVEING EKYCGERSQPVVTSNSNKITVRFHSDQSYTDT GFLAEYLSYDSSDPCPGQFTCRTGRCIRKELR CDGWADCTDHSDELNCSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPAQTF RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV
449	1799	A	3618	2	613	FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGITKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE
450	1800	A	3620	1	2676	MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDBADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKYLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE HRETYQKLLEDIAVLHRLAARLSSRAEVVGA

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	corresponding to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
<u></u>	ļ			sequence		nucleotide insertion VRQEKRMSKATEVMMQYVENLKRTYEKDH
						AELMEFKKLANQNSSRSCGPSEDGVLRTARS MSLTLGKNMPRRRVSVAVVPKFNALNLPGQ
		Ì				TPSSSSIPSLPALSESPNGKGSLPVTSALPALLE NGKTNGDPDCEASAPALTLSCLEELSQETKA RMEEEAYSKGFQEGLKKTKELQDLKEEEEEQ
						KSESPEEPEEVEETEEEKDPRSSKLEELVHFL QVMYPKLCQHWQVIWMMAAVMLVLTVVL
						GLYNSYNSCAEQADGPLGRSTCSAAQKDSW WSSGLQHEQPTEQ
451	1801	A	3623	504	198	QLIQHQTVHTGRKLYECKECGKAFNQGSTLI RHQRIHTGEKPYECKVCGKAFRVSSQLKQHQ RIHTGERPYOCKELKGRGAEMLAVLAVKEO
452	1802	Α	3628	2	195	NRTPVNYGK MTCLHSAKAFHY*SSCSFSCEEGFALIGPEVV
						QCTALGVWTAPAPVCIAVQCQHLEALNEGT MG*DYPFTAFAYGSSCKYECHTVYRVRGLD
				i		MLHSRGCYLWNGHFTT*EAISCEPLERPCH*S V*CSFSCEEGFALIGPEVVQCTALGVWTAPAP
453	1803	A	3637	662	142	VCIAVQCQHLEALNEGTMG IQAKGLGIWHVPNKSPMQHWR\KGSLLRYRT
						DTGFLQTLGHNLLGIYQKYPVKYGEGKCWT DNGPVIPVVYDFGDAQKTASYYSPYGQREFT AGFVQFRVFNNERAANALCAGMRVTGCNTE
	1					HHCIGGGGYFPEASPQQCGDFSGFDWSGYGT  WHVGYSSSREITE\AAVLLFYR
454	1804	Α	3641	1	362	TQVHPAMLGLDELGRSGCGHCTQADLRFGD AAGRDPGQDNDRNTAEPAFPPPPRVMAAAA
	i .	·				ALRAPAQSSVTFEDVAVNFSLEEWSLLNEAQ GCLYHDVMLETLTLISSLGKVLILNCDLS
455	1805	A	3646	2	414	AAAGRGASGALTGEGGGEQGRRVGLGSRAH SLLLGPTFNSCQVSSQPPRVAGLGLPLKHEPS
						RPQPPSPRGPRTVRAGVPGAHPQDTPCPEFVR PRKVPLVGEAPGLPPEERSRGWRRDTPGLQE
456	1806	A	3656	396	8	SRVRAPSYDDIT QIVSFNSYLTLYTKNNLKSMKDLNVNTEMIK
	i					LLELKNIHNLG*AKFFLN*IQKALIKRKILIHW P/LIKIK/SFCSLSDTIKKMKRQTIVWEQTFIIHI SVKELVSRIYEAFLQFNKTVNRPVFDIKKEQK
457	1807	Α	3660	14	1961	F SEAKLGGPTGMDLWQLLLTLALAGSSDAFSG
						SEATAAILSRAPWSLQSVNPGLKTNSSKEPKF TKCRSPERETFSCHWTDEVHHGTKNLGPIQLF
				·		YTRRNTQEWTQEWKECPDYVSAGENSCYFN SSFTSIWIPYCIKLTSNGGTVDEKCFSVDEIVQ
				. · [	Į	PDPPIALNWTLLNVSLTGIHADIQVRWEAPRN ADIQKGWMVLEYELQYKEVNETKWKMMDP
7.	•					ILTTSVPVYSLKVDKEYEVRVRSKQRNSGNY GEFSEVLYVTLPQMSQFTCEEDFYFPWLLIIIF
					. [	GIFGLTVMLFVFLFSKQQRIKMLILPPVPVPKI KGIDPDLLKEGKLEEVNTILAIHDSYKPEFHS
						DDSWVEFIELDIDEPDEKTEESDTDRLLSSDH EKLHINLGVKDGDSGRTSCCEPDILETDFNAH
						DIHEGTSEVAQPQRLKGEADLLCLDQKNQNN SPYHDACPATQQPSVIQAEKNKPQPLPTEGAE
						STHQAAHIQLSNPSSLSNIDFYAQVSDITPAGS VVLSPGQKNKAGMSQCDMHPEMVSLCQENF
						LMDNAYFCEADAKKCIPVAPHIKVESHIQP\S LNQEDIYTTESLT\TAAGSP\GTGEHVPGSEM

SEQID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- cotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	)	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	<u> </u>	}	}	amino acid	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
ļ	}			peptide	Soquence	/=possible nucleotide deletion, \=possible
		L		sequence	ł	nucleotide insertion
				]		PVPDYTSIHIVQSPQGLILNATALPLPDKEFLS
458	1808	A -	3663	154	462	SCGYVSTDQLNKIMP TRAPASGRSGAGLALSANAPDSGGHPGATEG
130	1000	^	3003	134	702	PAGSLAHASGSARGTWRVRGRGSHGWERTV
	ŀ	j	ļ	]	ļ	GAGGCANPVPALHSCASAPRGTGRVSALGPK
450	4000	<b></b> _				TGSSPLSSPKG
459	1809	Α	3664	902	135	LGKYNTSMALFDFVLHNSTGEIRYITEDDVIQ SQNALGKYNTSMALFESNSFEKTILESPYYVD
}		]	ļ		1	LNQTLFVQVSLHTSDPNLVVFLDTCRASPTSD
			İ			FASPTYDLIKSGCSRDETCK\VYPLFGHYGRF
		1	}		j	QFNAFKFLRSMSSVYLQCKVLICDSSDHQSRC
	}					NQGCVSRSKRDISSYKWKTDSIIGPIRLKRDR   SA\NGNSGFQHETHAEETPNQPFNSVHLFSFM
1			1			VLALNVVTVATITVRHFVNQRADYQ\YQKLQ
						NY
460	1810	A	3670	850	557	LGILMSPQVEAGEI*ALLTPPPGCMQFSPLTL/P
		1			[	K*WVSPGLTP/PPPEVPSVFLVEPGLPHAGQA GLDLL\TSGDPPASTSOSARTTDVSHRAOPLAI
1		Ì	ł			GLDLL(15GDFFAS15QSARI1DVSHRAQFLAI
461	1811	A	3671	2472	2099	IGVLAFETGSCSVTRLYCIGIIMPHCSLDLAGS\
1		[	[			TSAFRIAGTTSVHHHPQLTFFFFWIETGSHCV
]		l				VQTGL*LLALSNPPALASQIAGISGMSHRAWP GLVLYSLEFSLLCASQSLIMLFTCYNE
462	1812	A -	3672	394	110	VKPVNGESKRD•GADTQTCEGEADEQLQT\N
'	10.2	(	50,2	1	'''	CYYD/STKSFFYISCG*K\RKPTWAENRRLNA
		1				KMFGIPLHSNSDPWGYEEREVIGFHRSRVSRG
463	1813	A	3673	348	1	HGS QRNPFSAGHPQRPPTSGSQSELLAQPRLRPGR
103	1013	^	3073	346	1 *	KSSFSRDQDVW*SQAVPKRQ*QRNPFSAGHP
ĺ	•	ĺ	1 .	ĺ		QRPPTSGSQSELLAQPRLRPGRKSSFSRDQDV
ļ	ļ	1	Ì			WPGQKPRPSQQQHQMCASPTLGQRSPFALEP
			[			VPAYHGGRDPFASARPSPVGIPKPRAAPAGG GWRRIRPKSSTK
464	1814	A	3676	2253	320	PVIORCSOPYGFSLLISFFLKCVSETSOOPPSR
1		1				KVFQLLPSFPTLTRSKSHESQLGNRIDDVSSM
		1	1	]		RFDLSHGSPQMVRRDIGLSVTHRFSTKSWLS
	{	ì	1			QVCHVCQKSMIFGVKCKHCRLKCHNKCTKE APACRISFLPLTRLRRTESVPSDINNPVDRAAE
		1	1	ł	1	PHFGTLPKALTKKEHPPAMNHLDSSSNPSSTT
i			ţ	1	ĺ	FSTPSSPAPFPTSSNPSSATTPP\NPSP\GQR\DSR
	ļ	1				FNFPSC/AYFIHHR\Q\QFIFPDISAFAHAAPLPE
	.	ł	ì	ł	1	AADGTRLDDQPKADVLEAHEAEAEEPEAGK SEAEDDEDEVDDLPSSRRPWRGPISRKASQTS
1	Ì	ł	Ì	Ì	1	VYLQEWDIPFEQVELGEPIGQGRWGRVHRGR
	ł	1	ł	ì		WHGEVAIRLLEMDGHNQDHLKLFKKEVMN
			1			YRQTRHENVVLFMGACMNPPHLAIITSFCKG
		}	1.		ł	RTLHSFVRDPKTSLDINKTRQIAQEIIKGMGY LHAKGIVHKDLKSRNVFYDNG\KVVITDFGLF
	{	{	i			GISGVVPAEGRRENQLKLSHDWLCYLAPEIVR
1			1			EMTPGKDBDQLPFSKAADVYAFGTVWYBLQ
		1	ł	1	}	ARDWPLKNQAAEASIWQIGSGEGMKRVLTS
		1	1		1	VSLGKEVSENLSACWAFDLQERPS\FSLLMD
ł		}			1	MLEKLPKLNRRLSHPGHF*KSADINSSKVVPR     FERFGLGVLESSNPKM
465	1815	Ā	3679	8	803	IPSPAWWNSTWADTFSLLLALAVALYLGYY
	1	1				WACVLQTHRAFCASNTEDLETVVNHIKHRYP
l	1	1	}	1	l	QAPLLAVGISFGGILVLNHLAQARQAAGLVA
1	1	ļ		}	1	ALTLSACWDSFETTRSLETPLNSLLFNQPLTA GLCQLVERLSY/E*DLQARTIRQFDERYTSVA
L	1	J	ــــــــــــــــــــــــــــــــــــــ	l	L	GICCHARTISTIE DECARTIKALDER 112AV

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- cotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first amino acid	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ŀ	l	1	residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
1	ĺ	1	ł	peptide	1	/=possible nucleotide deletion, \=possible
	ļ	<u> </u>		sequence		nucleotide insertion
1			ļ			FGYQDCVTYYKAASPRTKIDAIRIPVLYLSAA DDPFSTVCALPKQAAQHSPYVALLITARGGHI
1						GFLEGLLPWQHWYMSRLLHQYAKAIFQDPE
166	1016	<u> </u>	2684		207	GLPDLRALLPSEDRNS
466	1816	A	3684	3	307	SSQYIVQSKTKIFL*AAREKQ/RHTCRRFSIRLS ANISSQTGEARGQWPSVFKVLKEKKLSTKKS
1		ł	]			FGQK*GR\RKTFPDKQK/LREFDTTRPTIQEML
162	1016	<u> </u>	2/97	04/5	02#	TGVLQG
467	1817	A	3687	2465	837	ELPTPLIAAHQLYNYVADHASSYHMKPLRMA RPGGPEHNEYALVSAWHSSGSYLDSEGLRHO
			[ [			DDFDVSLLVCHCAAPFEEQGEAERHVLRLQF
1			]			FVVLTSQRELFPRLTADMRRFRKPPRLPPEPE
1						APGSSAGSPGEASGLILAPGPAPLFPPLAAEVG MARARLAOLVRLAGGHCRRDTLWKRLFLLE
						PPGPDRLRLGGRLALAELEELLEAVHAKSIGD
l i				'		IDPQLDCFLSMTVSWYQSLIKVLLSRFPQSCR
}						HFQSPDLGTQYLVVLNQKFTDCFVLVFLDSH LGKTSLTVVFREPFPVOPODSESPPAOLVSTY
·		•		:		HHLESVINTACFTLWTRLL*GSGLDH*MSLFL
						ESWAYQIACQRQD*PALLGPRASQTLSDTKG
		,				FVTMS*GSAAPAWQQEPPSPNTHSH*PIQDSR ESGOPRGPLGPFWGTPFGPPGRVSGVHTGWO
	ł					TPPRAPLPESCPL\PLTTVSHLCPLSLRVFTSHL
					ļ	DITAGHSHRDDTWVPIPALPLKHLRPPSSPFA
	ľ			· ·		LGPWVSHPLMRWVQKLSHLHSNPGTGFSMG GKQQRN
468	1818	A	3691	960	499	QTCRKDKRAIYPHFQNE*MNEIKAI*SGTGGI
l i			ľ			QCFHSQNDSAFFFFLFLLETEFCSAA/TVQWH
] .]						DFLSMQPPPPGFKQFTCLSLLSSWNYRR\PPPF PGNF\*FLVKTGFPHVGQTGFELLTSSDLAPLA
						SQNGGITGMSPCAWPFFFFFFGLC
469	1819	A	3714	4747	495	MAYSWQTDPNPNESHEKQYEHQEFLFVNQP
	ĺ	ł	ĺ	Í		HSSSQVSLGFDQIVDEISGKIPHYESEIDENTFF VPTAPKWDSTGHSLNEAHOISLNEFTSKSREL
	ļ		.	`]	]	SWHQVSKAPAIGFSPSVLPKPQNTNKECSWG
		}	Ì	i		SPIGKHHGADDSRFSILAPSFTSLDKINLEKEL ENENHNYHIGFESSIPPTNSSFSSDFMPKEENK
		1	}		1	RSGHVNIVEPSLMLLKGSLQPGMWESTWOK
		ł	1			NIESIGCSIQLVEVPQSSNTSLASFCNKVKKIR
	ł	ł		ļ	1	ERYHAADVNFNSGKIWSTTTAFPYQLFSKTK FNIHIFIDNSTOPLHFMPCANYLVKDLIAEILH
(	1	- [	1		Ì	FCTNDQLLPKDHILSVWGSEEFLQNDHCLGS
		į	j	. ]	J	HKMFQKDKSVIQLHLQKSREAPGKLSRKHEE
	1			]	}	DHSQFYLNQLLEFMHIWKVSRQCLLTLIRKY DFHLKYLLKTOENVYNIEEVKKICSVLGCVE
	1			ļ	ļ	TKQITDAVNELSLILQRKGENFYQSSETSAKG
	]		- 1			LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV
	1		ľ	ļ		PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL
					l	AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM
[	- 1			}	\$	ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD
		·				SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW
		-		1	į	DERTVSEMHTILRRWTFSQPLEALGLLTSSFP
.	- 1	1		(		DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV
			}	1		KFEWNLESPLVQLLLHRSLQSIQVAHRLYWL LKNAENEAYFKSWYQKLLAALQFCAGKALN
		ŀ		1		DEPSKEQKLIKILGDIGERVKSASDHQRQEVL
	- 1	İ	1	ľ		KKEIGRLEEFFQDVNTCHLPLNPALCIKGIDH
						DACSYFTSNALPLKITFINANLMGKNISIIFKA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				·	_	GDDLRQDMLVLQLIQVMDNIWLQEGLDMQ MITYRCLSTGKDQRLVQMVPDAVTLAKIHRH SGLIGPLKENTIKKWFSQHNHLKADYEKALR NFFYSCAGWCVVTFILGVCDRHNDNIMLTKS GHMFHIDFGKFLGHAQTTFGGIKRDRAPFIFTS EMEYFITEGGKNPQHFQDFVVELCCRAYNIIR KHSQLLLNLL\EMMLYAG\PELSGI\QDLKY VYNNLRPQDTDLEATSHFTKKIKESLECFPVK LNNLIHTLAQMSAISPAKSTSQTFPQESCLLST TRSIERATILGFSKKSSNLYLIQVTHSNNETSL TEKSFEQFSKLHSQLQKQFASLTLPEFPHWW HLPPTNSDHRRFRDLNHYMEQILNVSHEVTN SDCVLSFFLSEAGQQTVEESSPVYLGEKFPDK KPKVQLVISYEDVKLTILVKHMKNIHLPDGSA
470	1820	A	3718	430	75	PSAHVEPYLLPYPSEVRRRKTKSVPKCTDPTY NEIVVYDEVTELQGHVLMLIVKSKTVFVGAI NIRLCSVPLDKEKWYPLGNSII*PLLLFSSFGM KSLEKDEFVGGMLLSNPIW SHGSISILNLHOGCVFLPSLPAOGLRCYRCLA
17,0	1020	A	3/16	450		VLEGASCSVVSCPFLDGVCVSQKVSV/CWQ*/ CPWGARAEGRLSAVVDSQISCCKGDLCNAV VLAAGSPWALCVOLLLSLGSVFLWALL
471	1821	A	3723	891	494	LRQSL/NSVPQAGVQWRDSSLQAPPPRFTPLS CLSLPSSWDYRRLPPCLANFLYF**RRGFTML ARMVLIS*PRDPPASASQ\STEITGGSHRAQHP TDSRDHSERSVKKSHEVISELRMKVIKCKVAF SKNPI
472	1822	A	3734	443	251	GFIET*NFCVSKDTSKKLS/RLPTKWKNVFAN *ISDKGLVSRICQELLRHLDAEQVSSTAGLSL
473	1823	A	3746	3	500	THASGARSGAGWAGRGVRAGTEAGRGGIF LTLSILRTRDLPSGAMSEGVDLIDIYADEEFNQ DPEFNNTDQIDLYDDVLTATSQPSDDRSSSTE PPPPVRQEPSPKPNNKTPAILYTYSGLRNRRA AVYVGSFSWWTTDQQLIQVIRSIGVYDVGEV KFAENRAK
474	1824	A	3753			RPLFAREGGIYAVLVCMQEYKTSVLVQQAG LAALKMLAVASSSEIPTFVTGRDSIHSLFDAQ MTREIFASIDSATRPGSESLLLTVPAAVILMIN TEGCSSAARNGLLLINLLLCNHHTLGDQIITQ ELRDTLFRHSGIAPRTEPMPTTRTILMMILNR YSEPPGSPAERAALETPIIQGQDGSPELLIRSLV GGPSAELLLDLERVLCREGSPGGAVRPLLKRL QQETQPFLLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL GKTCWEALVSPLVQNITSPDAEGVSALGWLL DQYLEQRETSRNPLSRAASFASRVRRLCHILL VHVEPPPGPSPEPSTRPFSKNSKGRDRSPAPSP VLPSSSLRNITQCWLSVVQEQVSRFLAAAWR APDFVPRYCKLYEHLQRAGSELFGPRAAFML

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid	to last amino acid residue of peptide	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
	_					ALRSGFSGALLQQSFLTAAHMSEQFARYIDQ QIQGGLIGGAPGVEMLGQLQRHLEPIMVLSG LELATTFEHFYQHYMADRLLSFGSSWLEGAV
						LEQIGLCFPNRLPQLMLQSLSTSEELQRQFHLF QLQRLDKLFLEQEDEEEKRL*EEEEEEEEEEA EKELFIEDPSPAISILVLSPRCWPVSPLCYLYHP RKCLPTEFCDALDRFSSFYSQSONHPVLDMG
						PHRRLQWTWLGRAELQFGKQILHVSTVQMW LLLKFNQTEEVSVETLLKDSDLSPELLLQALV PLTSGNGPLTLHEGQDFPHGGVLRLHEPGPQ
					·	RSGEALWLIPPQAYLNVEKDEGRTLEQKRNL LSCLLVRILKAHGEKGLHIDQLVCLVLEAWQ KGPNPPGTLGHTVAGGVACTSTDVLSCILHLL GQGYVKRRDDRPQILMYAAPEPMGPCRGQA
						DVPFCGSQSETSKPSPEAVATLASLQLPAGRT MSPQEVEGLMKQTVRQVQETLNLEPDVAQH LLAHSHWGAEQLLQSYSEDPEPLLLAAGLCV
						HQAQAVPVRPDHCPVCVSPLGCDDDLPSLCC MHYCCKSCWNEYLTTRIEQNLVLNCTCPIAD CPAQPTGAFIRAIVSSPEVISKYEKALLRGYVE SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK
						CGWASCFNCSFPEAHYPASCGHMSQWVDDG GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE KNEGCLHMTCAKCNHGFCWRCLKSWKPNH
						KDYYNCSAMVSKAARQEKRFQDYNERCIFH HQAREFAVNLRNRVSAIHEVPPPRSFTFLNDA CQGLEQARKVLAYACVYSFYSQDAEYMDVV EQQTENLELHTNALQILLEETLLRCRDLASSL
-						RLLRADCLSTGMELLRRIQERLLAILQHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEEDDEDDVPEWQQDEFDEELDNDSFS
475	1825	A	3754	1093	96	YDESENLDQETFFFGDEEEDEDEAYD GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ ATGRRRRTRTQQRTAALLTDGTTKTGAAW SPPDSLCWSSTTGABGAKAAU VIIS ATTTEN
						SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTR\RPASITGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA
		1		·		RSGGGRWRPNAFRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCIYCTNHVSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR
476	1826	A	3758	901	521	PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW\DYRHVPPRQANFCIF/M*RRG
		·				FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI
				843		GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTIPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL
478	1828	A	3763	267		HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY
						WTGHSFASQAWLRQVPEVSKHLQCPSAESLL TMEYHQPEDPAPGKAGTAEAVIPENHEVLAG PDEHPQDTDARDADGEAREREP/RRPSFAA*P
				·		VWGQPESPLPEASSAPPGPTLGTLPEVETIRA CSMPQELP*SPRTRQPEPDFYCVKWIPWKGE QTPIITQSTNGPLPSPCHHEHPLSSVEGEAPPA

WO 01/57188

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ĺ	ĺ	914	ng to first	acid residue	Q=Glutzmine, R=Arginine, S=Serine,
1 .			ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
'		ľ	i	peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
		<u> </u>		Sequence		EGSDHIG
479	1829	A	3766	2	2152	YSPIRLLEVCVPLPKIFIKRQAPLKVSLLQDLK
1		'		_		DFFQKVSQVYVAIDERLASLKTDTFSKTREEK
i i			1			MEDIFAQKEMEEGEFKNWIEKMQARLMSSS
						VDTPQQLQSVFESLIAKKQSLCEVLQAWNNR
						LQDLFQQEKGRKRPSVPPSPGRLRQGEESKIS
						AMDASPRNISPGLQNGEKEDRFLTTLSSQSST
						SSTHLQLPTPPEVMSEQSVGGPPELDTASSSE
1						DVFDGHLLGSTDSQVKEKSTMKAIFANLLPG
1						NSYNPIPFPFDPDKHYLMYEHERVPIAVCEKE PSSIIAFALSCKEYRNALEELSKATQWNSAEE
1						GLPTNSTSDSRPKSSSPIRLPEMSGGQTNRTTE
1						TEPOPTKKASGMLSFFRGTAGKSPDLSSOKRE
, ·						TLRGADSAYYQVGQTGKEGTENQGVEPQDE
						VDGGDTQKKQLINPHVELQFSDANAKFYCRL
1 1						YYAGEFHKMREVILDSSEEDFIRSLSHSSPWQ
						ARGGKSGAAFYATEDDRFILKQMPRLEVQSF
						LDFAPHYFNYITNAVQQKRPTALAKILGVYRI
						GYKNSQNNTEKKLDLLVMENLFYGRKMAQ VFDLKGSLRNRNVKTDTGKESCDVVLLDENL
1 1						LKMVRDNPLYIRSHSKAVLRTSIHSDSHFLSS
}						HLIIDYSLLVGRDDTSNELVVGIIDYIRTFTWD
						KKLEMVVKSTGILGGQG*MPTVVSPELYRTR
	4000					FCEAMDNYFLMVPDHCTGLGLNC
480	1830	A	3777	251	3	QGCGSAGTLIHY**ECKMVQLLWKTV*QFLI
		I				KLNI/KDPAITLDVYPNEVKNYVRTKTYTQMF
481	1831	A	3779	333	3	I/ANFIMAKSWKQPTHPSVRT EAAIRQPEPNILDVNQIFKDLAMIIHDOGDLID
"	103.	^.	3,,,	333	3	SIEANAESSEVLVERAPGQLQRPA\YYQKKSR
	ļ	- 1		ļ	,	KKMCLVVLVQTAIILICERIM*VVYTTKWSPPI
		[				VLPVSCFQGQKFN
482	1832	A	3780	2	371	TGGRQGKNDHTSITEKPSRDFNRHLITQNI*M
1 1	1	ł	ł	ł		PNQDMKSSSNSLIIRKVQIKPTILYHHIFTRKA
			l			KMKTTDKTKYR*GFKAITTLIHCSQDCKLQ*S
483	1833	A	3787	43	440	/L*ENHFMIFPKAEQHITYDTTIPFLR
703	1033	^	3/8/	43	448	LMKDLSPYVMETHYILNRLNER/RSMWRHIIG
	}	- I	j	,		KLPNTKDQEKILKAIRGRREVIQGS/RQQYRR PAAFSAAEKARRLWCS/VFNIERRNL/CEYPTK
	ľ	ı			·	LSFNIKGEMTFSDKTEFTTNRPSLKMLLKDRI
					1	QEEGKMF*KEKCFKRKE
484	1834	A	3798	1	727	FFFFETESRSVAQAGVQWCNLGSLQALPPGF\
]	J	J	)	ļ	. ]	SHSPASASRVAGTTGTRH*ARLIFYIFSRDGVS
					1	PC*PGWS*SPDLVIRPP\RLPKCWDYRREPPRP
	l	l			!	A*FFVFLVE\QGFTMLARMVSIS*PQ/CDLPAS
	j	. [	1	Į	j	VSQNAGITGVSHCAWPCLHFCFFGFFFEMESC
	į	]	Į	1	Ì	SVAQAEVQWHDLRSLQAPPPGFTPFSCLSLPG SWDYRRPPPRPANF\CIFSRDGVSPC*PGWSRS
	l			į	ŀ	PDLVIRPPRPPKVLGLQA
485	1835	A	3802	1	239	FFFFEMECLTVSQAGVQWYNLHSLQPLPPGF
	,	-		- J		KQFSC\LSLPSSWD*RVPTSRPAKF/CVIF*DGV
						SHCQPGWSAVVQPPLH
486	1836	Α	3811	378	98	RYD*SSQSENIP\QKEFLLKYP*CTATLGMRN
1			į	ļ	l	MSIMKKKSIFSAEFYKVSLPSLLL\HLLAIEWG
407	1000					FHIEIQLTIHQHFLNYELESDFVHIVEYM
487	1837	A	3814	771	320	FDPDWTRAAGIRHEKKPKALAYRRENSPGDL
	İ	1				PPPPLPPPEEEASWAL/GAEGSRQHVLPGAGA
		- 1			.	QWGEESGPGRAPGSPAGAPPR*RGLAPNSRP
		ĺ			l	SFLSRGQGTSTCSTAGSNSSRGSSSSRGSRGPG
				l		RSRSRSQSRSQSQRPGQKRREEPR

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO:	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	seq- uence		USSN 09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
488	1838	A	3818	sequence 1	781	nucleotide insertion FRACILELIPYAPILSWTACPPAMAGPRGLLP
	·					LCLLAPCLAGFSFVRGQVLFKGCDVKTTFVT HVPCTSCAAIKKQTCPSGWLRELPDQITQDCR
						YEVQLGGSMVSMSGCRRKCRKQVVQKACCP GYWGSRCHECPGGAETPCNGHGTCLDGMDR
		!				NGTCVCQENFRGSACQECQDPNRFGPDCQSV CSCVHGVCNHGPRGDGSCLCFAGYTGPHCD
400	1820					QELPVWQELGFPQNNPRLRKAPNCKCLPG*H RNGLIATPNPCRP
489	1839	Α	3822	934	669	FFFSEMESRSVTRLECSGAISAHLRLLGSSNSP ASAS*VAGTIGACHHAQLIFVFLVETGFHHVG
490	1840	Α	3825	79	9748	QDGLDLL/NLMIHPPRPPKVLGFQA GCQSCWPAWPRLRRRGPASAGARLGRKAPW
						GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL
						KSFQQQQQQQQQQQQQQQQQPPPP PPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPPP
						GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA
						ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK
						CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI
						RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV
						KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL
					ļ	LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS
						ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDITTEQPRSQHTLQADSVDLASCDLTSS
						ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD
						GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF
						VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV
		l			*	SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEEQYVSDILNYIDHGDPQVRGATAILC
		Ī				GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM
						SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEL LBTLAEIDFRLVSFLEAKAENLHRGAHHYTGL
		İ				LKLQERVLNNVVIHLLGDEDPRVRHVAAASL
				]		IRLVPKLFYKCDQGQADPVVAVARDQSSVYL KLLMHBTQPPSHFSVSTITRIYRGYNLLPSITD
		ŀ				VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDESR
		1				KSCTVGMATMILTILLSSAWFPLDLSAHQDAL ILAGNLLAASAPKSLRSSWASEEEANPAATK
			ļ		•	QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK
		1				GKEKEPGEQASVPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA
		İ			į	NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGKCVEEILGYLKSCFSREPMMATVC
						VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA
					ľ	SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAIHNHIRLFEPLVIKALKQ
						The state of the s

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, I=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
					·	YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW
						AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY
	÷.					LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRINTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATILGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD
	·				*	PVPSLSPATTGALISHEKLLLQINPERELGSMS YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL LELYSRWILPSSSARRTPAILISEVVRSLLVVS DLFTERNQFELMYVTLTELRRVHPSEDEILAQ YLVPATCKAAAVLGMDKAVAEPVSRLLESTL RSSHLPSRVGALHGVLYVLECDLLDDTAKQL IPVISDYLLSNLKGIAHCVNIHSQQHVLVMCA TAFYLIENYPLDVGPEFSASIIQMCGVMLSGS EESTPSIIYHCALRGLERLLLSEQLSRLDAESL VKLSVDRVNVHSPHRAMAALGLMLTCMYT GKEKVSPGRTSDPNPAAPDSESVIVAMERVS VLFDRIRKGFPCEARVVARILPQFLDDFFPPQ DIMNKVIGEFLSNQQPYPQFMATVVYKVFQT LHSTGQSSMVRDWVMLSLSNFTQRAPVAMA TWSLSCFFVSASTSPWVAAILPHVISRMGKLE QVDVNLFCLVATDFYRHQIEEELDRRAFQSV
491	1841	A	3826	469	302	LEVVAAPGSPYHRLLTCLRNVHKVTTC SNPPASASRVAGITGVHQHAWLIFVFLVEMEF
492	1842	A	3836	392	88	HHVGQAVLKLLISGDLPVSASQSA VAPSPMIMPDLYFYRDPEEIEKEE*AAAEK\EE FQSEWTAVV/PIEFTATQSEVADWFKDMQVP SVPIQQFPTEDWST*PTMNDWSATSTAQTTE WVRITTEWP

NO. of   Not.	SEQ ID	SEQ ID	Met	SEQ	Predicted	Decidend and	TA-t
Decide   Deptide   Decide   Dissay   Decide   Dissay   Decide   Dissay   Decide   Dissay   Decide   Dissay   Decide   Dissay   Decide   Dissay   Decide   Dissay   Decide   Dissay   Decide   Dissay   Decide   Dissay   Decide   Dissay   Decide   Dissay   Decide   Dissay   Decide						Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence			nou	1 -			
	1	,		1 —			
gen first amino acid residue of peptide residue of peptide pequence   Q-Glutsmine, R-Arginine, S-Serine, Carpine, W-Trypropinen, V-Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide delicion, \possible sequence   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide delicion, \possible nucleotide insertion   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide delicion, \possible nucleotide insertion   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide delicion, \possible nucleotide insertion   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide delicion, \possible nucleotide insertion   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide insertion   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide insertion   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide insertion   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide insertion   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide insertion   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide insertion   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide insertion   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide insertion   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide insertion   Tryprosine, X-Unknown, №50p codon, Y-Sposible nucleotide insertion   Tryprosine, X-Unknown, P-Sposible nucleotide insertion   Tryprosine, X-Unknown, P-Sposible nucleotide insertion   Tryprosine, X-Unknown, P-Sposible nucleotide insertion   Tryprosine, X-Unknown, P-Sposible nucleotide insertion   Tryprosine, X-Unknown, P-Sposible nucleotide insertion   Tryprosine, X-Unknown, P-Sposible nucleotide insertion   Tryprosine, X-Unknown, P-Sposible nucleotide insertion   Tryprosine, X-Unknown, P-Sposible nucleotide insertion   Tryprosine, X-Unknown, P-Sposible nucleotide insertion   Tryprosine, X-Unknown, P-Sposible nucleotide insertion   Tryprosine, X-Unknown, P-Sposible nucleotide insertion   Tryprosine, X-Unknown, P-Sposible nucleotide insertion   Tryprosine, X-Unknown, P-Sposible nucleotide inser			1				l=Isoleucine, K=Lysme, L=Leucine,
### ### ### ### ### ##################		ucia	1			1	
Persidue of peptide   Sequence   Persidue of peptide   Sequence   Persidue uncloid delicion, Impossible   Sequence   Sequence   Sequence   Persidue uncloid delicion, Impossible   Sequence   Sequence   Persidue uncloid delicion, Impossible   Sequence   Persidue   Persidue uncloid delicion, Impossible   Sequence   Persidue   Pe	delice	į.	1	914			Q=Glutamine, R=Arginine, S=Serine,
1843	1	l .	ı				1=Threonine, V=Valine, W=Tryptophan,
1843   A   3838   19   380   TISSIMANAPETITOSIGEENASSPSPIPYFERK   KRKRS-PRAHRYKIDOPEDIPI KUPI-CKHSK   KRKRS-PRAHRYKIDOPEDIPI KUPI-CKHSK   KRKRS-PRAHRYKIDOPEDIPI KUPI-CKHSK   KRKRS-PRAHRYKIDOPEDIPI KUPI-CKHSK   KRKRS-PRAHRYKIDOPEDIPI KUPI-CKHSK   KRKRS-PRAHRYKIDOPEDIPI KUPI-CKHSK   KRKRS-PRAHRYKIDOPEDIPI KUPI-CKHSK   KRKRS-PRAHRYKIDOPEDIPI KUPI-CKHSK   KRKRS-PRAHRYKIDOPEDIPI KUPI-CKHSK   KRKRS-PRAHRYKIDOPEDIPI KUPI-CKHSK   KRKRS-PRAHRYKIDOPEDIPI KUPI-CKHSK   KRKRS-PRAHRYKIDOPEDIPI KUPI-CKHSK   KRKRS-PRAHRYKIDOPEDIPI KUPI-CKHSK   KRKRS-PRAHFY KUPI-LKHSK   KRSIS-PROJE KRSIS-PR	J	j	Į.			sequence	Y=1 yrosine, X=Unknown, *=Stop codon,
1843   A   3838   19   380	İ		ļ				/=possible nucleotide deletion, \=possible
1844   A   3845   2   352	402	1040	<del>                                     </del>	2000	<u> </u>	<del> </del>	
1844   A   3845   2   352   FFFLRRSLDSVAQARAQWIGEGLIQAPPPIA	493	1843	A	3838	19	380	
1844   A   3845   2   352	1	ľ	l	1	·		
1844   A   3845   2   352   FFFERRSLDSVAGDAGQWAELGLLQAPPER		ŀ	Ì				
SPIS_PAGL_PSSWD_TGRPPPC2ANF_CIP/MPRRG		<del> </del>	<u> </u>				
### 1845 A 3847 1774 40 DIFFREAKEGNGQDEAQFSVEMPLTGKAYL ### 40 DIFFREAKEGNGQDEAQFSVEMPLTGKAYL ### ARPODIDELY AHQGRCWFRIL ### 40 DIFFREAKEGNGQDEAQFSVEMPLTGKAYL ### ARPODIDELY AHQGRCWFRIL ### ARPODIDELY AHQGRCWFRIL ### DEPARTMENT OF THE WARKYNQTHY DEPARTMENT OF THE WARKYNQTHY DEPARTMENT OF THE WARKYNQTHY DEPARTMENT OF THE WARKYNQTHY DEPARTMENT OF THE WARKYNQTHY BEACADINKDFALLRHAGPFYEDLAFKKYNTEW EYSHEHIGRCGPANGFGQ WHERKYRYTEW RPWGTAGRCFRGHSKGASVKLVVTPGPLSGL GORGTISHLERHISARAVGPHFKGRGHH*AC HGELRRHWDRLA*GPDATEGALGASFEIEG GORPADLTVQADTLHEPSARLGGAHRACPK RRPHRVLWRWARGAWAWRCQAREKQETQG QPCHITGHPLGRBAEPAAAGALARRPFF ARTGSTBWGPCWRPIRHCRODDLWPTPTCARD WPPTIFPVLAGGHPADALARRPFF ARTGSTBWGPCWRPIRHCRODDLWPTPSFAHR TVASHSPFSGQGBGRGPHGCESPGRSGPAGR LVLQHPTGTSFTAAKKVRPGDPLTSPFATSTRAT TVASHSPFSGQGBGRGPHGCESPGRSGPAGR LVLQHPTGTSFTAAKKVRPGPEPTTSPVT SPRPTAPPRHP ASSGNSVCFSKKTCRWEKK SPVLMELAYWQDAWT SPRPTAPRHP ASSGNSVCFSKKTCRWEKK SPVLMELAYWQDAWT LSWDYXSLFPRPNYCLLVELGFHHVQAG LKLLTSSADPALASQSABITGSHRWFVELLIA RRPVIRIRAPPQRLFFNLTISLKALSFPMATT  ### 1847 A 3859 2 393 ALRETTREDISARGAGGAGSHRWFVELLIA RRPVIRIRAPPQRLFFNLTISLKALSFPMATT  ### 1847 A 3859 2 393 ALRETTREDISARGAGGAGGAGFTPGCHLFVPR RPPAKKGLPSDTPHSKAPTPHLILGGEDSQ VPIL  ### 1848 A 3860 253 634 KNASTVYSSQCDPKSFFILRWSLALVAQAG EQ*RODLSSLQPPPGRYGPG PEGAGPSPPPGTPGRGGGSSSEGPPQCLLFVPR RPPAFKKGLPSDTPHSKAPTPHLILGGEDSQ VPIL  ### 1849 A 3863 423 663 ARSGNIVSSRAWPRHIPLYWKTTPL  ### 4P\$ 1849 A 3863 423 663 ARSGNIVSSRAWPRHIPLYWKTTPL  ### 4P\$ 1849 A 3863 423 663 ARSGNIVSSRAWPRHIPLYWKTTPL  ### 4P\$ 1849 A 3863 423 663 ARSGNIVSSRAWPRHIPLYWKTTPL  ### 4P\$ 1849 A 3865 2 15246 PREGCLWCLAGRSFTARFOPSRPARSPLPLFP DILRWASDLDIMGDAGGBGEVQFLRTDDEV VLQCATVLKEQKIKLCLAAGEGVESTRAUPDISNISVRNVIGI LVGIALNI.  ### 4P\$ 1849 A 3865 2 15246 PREGCLWCLAGRSFTARFOPSRPARSPLPLFP DILRWASDLDIMGDAGGGEVQFLRTDDEV VLQCATVLKEQKLCLAAGEGVALFRANGTLIRIHAN ANTYSAGVESSGGWAVTHASDRGTISHERADS DDQRRLVYTYRGGAVCTHASSLQVALERANGTLIRIHAN ANTYSAGVESSGGWAVTHASDRGTISHERADS DDQRRLVYTYRGGAVCTHASSLQVALERANGTLIRIHAN SRM1SCLTTRSSMTDALIPHOLCCLTSPADS DDQRRLVYTYRGGAVCTHASSLQVALE	494	1844	Ι Α	3845	2	352	
ARRODIDELYAHQGRCWFRLL			1	1	ĺ	(.	
1845   A   3847   1774   40	!		[	1	ł		FTVLARMVLIS*PCDPPTLASQGTAITGMSYH
### WADKYRPRZPRTPRIVHIGEEWNKYNOTHY   DEPNPPPKIVQGYKPNUYDLDKKSTPEYFL     EACADNKDP ALIRHAGPPYEDIAFKIVNREW     EYSHRHIGFRCQFANGIFQLWFHIPKYRYRR*     RPWGTAGRCFGHSKGASVLVVTPOPLSGL     QGRGTTSHLRPHLSPRARQGAFKRGETOG     QGRGTTSHLRPHLSPRARQGAFKRGETOG     QQPADLTVQADTLHRPSARLGGAFRACPK     RPPHRVLWR WARGAW WAYGCAREKGETOG     QPCHTTGHPLGRRAEPAAGAAPALAHPPF     ARTGSTEWQPCWRPHACRSDPLWTPTLCRD     WPPTHPVLAGGVHFPAAGAGGCVEVPVSVN     WMGTKSH*AVLPPPSTGPGQGLPBGWGLE     KGEGLPGUFPPGLLTGPWSMCPTSHARR     TVAFSHSPRSGQBGRGPHGCHSPGRNSGPAGR     LVLQHPTGTSPTBARRKVPPGFHFTSPTAH     TVAFSHSPRSGQBGRGPHGCHSPGRNSGPAGR     LVLQHPTGTSPTBARRKVPGFHFTSPTAH     TVAFSHSPRSGQBGRGPHGCHSPGRNSGPAGR     LVLQHPTGTSPTBARKKVPGFHFTSPTAH     TVAFSHSPRSGQBGRGPHGCHSPGRNSGPAGR     LVLQHPTGTSPTBARKKVPGFHFTSPAFT     TVAFSHSPRSGQBGRGPHGCHSPGRNSGPAGR     LVLQHPTGTSPTBARKKVPGFHFTSPAFT     TVAFSHSPRSGQBGRGNSVCTSKKTCRWEK     SVVLMELAYWQDRMFF     SRPPTAPRHPASSGNSVCTSKKTCRWEK     SVVLMELAYWQDRMFF     SRPPTAPRHPASSGNSVCTSKKTCRWEK     SVVLMELAYWQDRMFF     SRPPTAPRHPASSGNSVCTSKKTCRWEK     SPVLMELAYWQDRMFF     LLSWMVTSLFPPNTGLLVLGFHHVQQAGR     LLLTSSALPALASQSABTIGMSHRWPLPLL     RPPVRIRAPPGLPPNTTISLKALSPSMATT     A 3859 2 393 ALRKTRRDGIARTGAGPASSWKGTNNYPWR     LEMAGRPGQSQCSKDRGTGSLPPSQRFLGF     PEGAGAPSPPPPGTRGGGSSSEGPPQLLFVPR     RPPAFKKGLPSDTPHSKAPTTHLLGGEDSQ     PURL     LEMAGRPGQSQCSKDRGTGSLPPSQRFLGF     PEGAGAPSSPPGTRFRCHGGSSSEGPPQLLFVPR     RPPAFKKGLPSDTPHSKAPTTHLLGGEDSQ     PURL     498 1848 A 3860 253 634 KNASTVYSQGDPKSFFLLRWSLAVAQAG     PURL		<u> </u>		<u> </u>			ARPQDIDFLYAHQGRCWFRLL
WADKYRPRKPRFPRVRUTGFEWNKYNOTHY    DFDNPPFKVQOYKPRIPDLIDKRSTPEYFL     EACADNKDFAILRHIAGPPYEDLIDKRSTPEYFL     EACADNKDFAILRHIAGPPYEDLIDKRSTPEYFL     EACADNKDFAILRHIAGPPYEDLIDKRSTPEYFL     EACADNKDFAILRHIAGPPYEDLIDKRSTPEYFL     EACADNKDFAILRHIAGPPYEDLIDKRSTPEYFL     EACADNKDFAILRHIAGPPYEDLIDKRSTPEYFL     EACADNKDFAILRHIAGPPYEDLIDKRSTPEYFL     EACADNKDFAILRHIAGPPYEDLIDKRSTPEYFL     EACADNKDFAILRHIAGPPYEDLIDKRSTPEYFL     EACADNKDFAILRHIAGPPYEDLIDKRSTPEYFL     EACADNKDFAILRHIAGPPYEDLIDKRSTPEYFL     EACADNKDFAILRHIAGPPYEDLIDKRSTPEYFL     EACADNKDFAILRHIAGPPYEDLIDKRSTPEYFL     EACAGRPGOGLEPOCHYPSVN     VMGTKSIF AVLPPPSTGOGGCSPDFWGLE     KGGLPFGIPPFOLLIGPPWSMRPPYTSPAHRR     TVAPHSIPSFOQGCRSPPSQGLPEGMGLE     KGGLPFGIPPFOLLIGPPWSMRPPYTSPAHRR     TVAPHSIPSFOQGERSPROGGLPSGWGLE     KGGLPFGIPPFOLLIGPPSCPATCR     VAPHSIPSFOQGERSPROGGLSFAILRHIAGPPYEDLITSPVT     SPRPTA AVLPPPSTGOGGSPGRSKKRLTCRWEKK     SVVLMELA Y WODKMFF     LISSNDYNSLSPPRVINGLEVELGFHHYDQAG     LILLISSALPALASGSAEITGMSHRIWPILLIK     RPPVIRRAPPORLFFILLISKLAISPMMATF     497	495	1845	Α	3847	1774	40	DIFFRRAKEGMGQDEAQFSVEMPLTGKAYL
DEPNIPPEKIVQOYKENIFYPDLIDKRSTPEYFL   EACADNINGPAILRPHIAGPEDIAFKIVNREW   EYSHRHGFRCQFANGIFG]   WYFHFREYRYRR'R   RPWGTAGRCRGHISKGASVKLVVTPGPLSGL   QGRGFTSHLRPHLSFARPQPPIPKGGHH'AC   HGELRRHWDRLA'GPDATEGALGASFEHEG   GQQPFADLITQADTILHRSALGGAHRACPK   RRHRVLWRWARGAWAWRCQAERKQETQG   QPCHTIGHIPGREAEPAAGAPALARIPPF   ARTGSTEPGPCWRPIRHCREDPLYTPTIC.CRD   WPFTHPVLAGGWHFPAAGAPALARIPPF   ARTGSTEPGGPCWRPIRHCREDPLYTPTIC.CRD   WPFTHPVLAGGWHFPAAGAPALARIPPF   ARTGSTEPGGPCWRPIRHCREDPLYTPTIC.CRD   WFFTHPVLAGGWHFPAAGAPALARIPPF   ARTGSTEPGGECGPHYGLITGPWSMRPVTPSFAHIR   TVAFSHSPFSGQEGRGPHCGHSFGRUGEVAGG   LVLQHFTGTSTTEAKRKVPFGPFEGHTSPVT   SPRPFTAPPHPASGNISSVCFSKKTCRWEKK   SFVLMELAYWQDRMFF   STRPFTAPPHPASGNISSVCFSKKTCRWEKK   SFVLMELAYWQDRMFF   STRPFTAPPHPASGNISSVCFSKKTCRWEKK   SFVLMELAYWQDRMFF   STRPFTAPPHPASGNISSVCFSKKTCRWEKK   SFVLMELAYWQDRMFF   STRPFTAPPHPASGNISSVCFSKTCRWEKK   SFVLMELAYWQDRMFF   STRPFTAPPHPASGNISSVCFSKTCRWEKK   SFVLMELAYWQDRMFF   STRPFTAPPNICTURGHTSPLVL   RRPFVRRRAPGRLFFNLITSLKALSPNMATT   LSMGSPLFSCFFLFNLITSLKALSPNMATF   LEMGAGRISGEPTNITSLKALSPNMATF   LEMGAGRISGEPTNITSLKALSPNMATF   LEMGAGRISGEPTSCRIPPSGPFLGPS   PEGAGPSPFPFGREGGSSSSSSFPPLGEPSQPFLGPS   PEGAGPSPFPFGREGGSSSSSSFPPLGEPSQPFLGPS   PEGAGPSPFPFGREGGSSSSSSFPPLGEPSQPFLGPS   PEGAGPSPFPFGREGGSSSSSSFPPLGEPSQPFLGPS   PEGAGPSPFPFGREGGSSSSSSFPPLGEPSQPFLGPS   PEGAGPSPFPFGREGGSSSSSSFPPLGEPSQPFLGPS   PEGAGPSPFPFGREGGSSSSSSSFPPLGEPSQPFLGPSQPFLGPS   PEGAGPSPFPFGREGGSSSSSSSFPPLGEPSQPFLGPSQPFLGPSQPFRAPPQFRFRAPTGNAPTKIP   MTSTFTEERSCHERTMANDFTLTSLKALSWALDELALARGFGNRLCFFLEP   TSNAGNVPPDLAICCFVLEQSLSVRALQEML   ANTYPAGVESSNGGGHRYTHMSFWGDDILVSVS   SERYLHLSTASGELQVDASFMQTLWNMPPC   SRCEEGFVTGGHVLRJFHGHMDELJTSPADS   DDCRRLVYYEGGAVCTHARSLWRLEPIRIS   WSGSILRWGQFLVRRYTTGGHLALTEDQG   LVVVDASSAMTKATSFFCFTSISKEKLDLATEDQG   LVVVDASSAMTKATSFFCFTSISKEKLDLATEDQG   LVVVDASSAMTKATSFFCFTSISKEKLDLATEDQG   LVVVDASSAMTKATSFFCFTSISKEKLDLATENG   LVVVDASSAMTKATSFFCFTSISKEKLDLATENG   LVVVDASSAMTKATSFFCFTSISKEKLDLATENG   LVVVDASSAMTKATSFFCFTSISKEKLDLATENG   LVVVDASSAMTKATSFFCFTSISKEKLDLATENG   LVVVDASSAM	i	ľ					WADKYRPRKPRFFNRVHTGFEWNKYNOTHY
EACADNKDFALLR/HAGPYYEDIAKKIVNREW EYSHRIGIFRQVANGIFICAQUANIFIKRY XYRR RPWGTAGRCPRGIBKGASVKLVVTPGPLSGL QGRGFTSHLRPHLSFARPQFPPI*VGGHHYAC HGELRRHWDRLA*GPDATEGALGASFEREG GQPPADLTVQADTLHRFSARLQGCVEVPVSVN RRPHRVLWRWARGAWAWRCQAREKQETQG QPCHITGIPLGREAEPAAGAPALARRPF ARTGSTEPPGPCWPRIPRCEPPLWTPTLCRD WPPTHPVLAGGVIEPPAAGIGGCVEVPVSVN VMGTKSH*AVLPPPPSTGPGQGLPEGWGLE KGBGLPPGIPPGLLTGRYMRPVTPSFAHR TVAFSHSFSGQEGRGPHGCHSGRSGPAGR LVLQHTGTSTTEAKRKVPGPPEGHFTSPVT SPRPTAPPRHPASSGNSSVCFSKKTCRWEKK SFVLMELAYWQDRMFP  496 1846 A 3849 830 442 AKSPLPLG*HQWRANLGSLKLRLPGFK*FTCLQ LLSSWDYRSLPPRVNFCILVELGFHHVDQAG LKLLTSSALPALASGASEIGMSHRIWPLPLLR RPPVRRAPPQRLPFNLTTSLKALSFMATF 497 1847 A 3859 2 393 ALRKTRRDGJARTGAQPASWKGTNNYPWR LEMAGRPGSQEGSKDRGTGSLPPSGRPLGPS PEGAGPSPPPGIPRGGASSEGPPQLLFVPR RPPAFKGLPSDTPHSKAPPTHLLIGGEDSQ VPIL LEMAGRPGSQEGSKDRGTGSLPPSGRPLGPS PEGAGPSPPFGFRFGRSGSESGPPQLLFVPR RPPAFKGLPSDTPHSKAPPTHLLIGGEDSQ VPIL 498 1848 A 3860 253 634 KNASTVYSSQGDFKSFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGFK*FSCLSLPSSWDYNCG- LPCLANPY*LVETGHHIVGQADLKLITSGDP PTSASESAGTTGVSTRAWRHIFLYWKTFFL 499 1849 A 3863 423 263 APSGISVATLYBRITTHMKEIK/DTDKW KDILY*WTGKIN*KMSTPFKATYRFNATYKLP MTFTEEKSIKIKTWNHKKPPNTQSNECKE*S FCSLLWYFGGFLWFMATPMPDFSISVKNVIGI LVGIAINL ANTYFEAGVESSGGGMRLCFLEP TSNAGNVPPDLAICCFVLEQSLSVRALQEML ANTYFEAGVESSGGGGGTRLJCFHLEP TSNAGNVPPDLAICCFVLEQSLSVRALQEML ANTYFEAGVESSGGGGGTRLJCFHLEP TSNAGNVPPDLAICCFVLEQSLSVRALQEML ANTYFEAGVESSGGGGGTRLJCFHLEP TSNAGNVPPDLAICCFVLEQSLSVRALQEML ANTYFEAGVESSGGGGGTRLJCGHLJVSNS SERYLH.STAAGELQVDASFMQTLWNMMPPC SRCEEGFVTGGHVLR.HFIGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPIRIS WSGSILRWGQFLXVRHVTTGGVLAITEDQG LVVVDASKAATHKATSFCFTGRISKEKLDLATEDQG LVVVDASKAATHKATSFCFTGRISKEKLDLATEDQG LVVVDASKAATHKATSFCFTGRISKEKLDLATEDQA	ł	ļ	ļ	1	ļ		DFDNPPPKIVQGYKFNIFYPDLIDKRSTPEYFL
RPWGTAGRCPRGHSKGASVKLVTYGELGEL		ĺ	l	-			EACADNKDFAILRFHAGPPYEDIAFKIVNREW
RPWGTAGRCPRGHSKGASVKLVTYGELGEL	1	1	ļ	1			EYSHRHGFRCQFANGIFQLWFHFKRYRYRR*
GGRGFTSHLRPHLSFARPQFPPFKGGHH*AC     GGLRRHWDRIA.4GPTGGLIGASFEBEG     GQQPADLTVQADTLHRPSARLGGAHRACPK     RRPHRVLWWARGAWAWRCQAEKGETGC     QPCHITIGHPLGRAEBAAGAAPALAHRPPF     ARTIGSTBPGPCWRPRHCRRDPLWTPTLCRD     WPPTHPVLAGGVHPPHAGIGGCVEVPVSVN     WMGTKSH*AVLPPPSTGPGGGGLPGWGLE     KGGGLPPGIPPGLILTGPWSMRPVTPSPAHIR     TVAPSHSPFSGQERGPHGCHSPGRKSGPAGR     LVLQHPTGTSPTIEAKRKVPPGPPGHPTSPVT     SPRPTAPPHPASSISSNSVCFSKKTCRWEKK     SFVLMELAYWQDRMFF     496	)	l	j	J			RPWGTAGRCPRGHSKGASVKLVVTPGPLSGL
HGELRRIHVDRI.A*GPDATEGALGASFEHEG	1						QGRGFTSHLRPHLSFARPQFPPI*KGGHH*AC
GQQPPADLTVQADTLHRPSARLGGAHRACPK   RPPIRRVLWRWARGAWARCQAREKOETXG   QPCHITGHPLGREAEPAAAGAAPALAHRPPF   ARTGSTBYGDPCWRPIRHCRRPPL WITTLCRD   WPPITHPVLAGGWIEPFAAG/IGGCVEVPVSVN    WMGTKSH*AVLPPPPSTOPGQGLEGWGLE   KGGGLPFGIPPFOLLTGPWSMRPPYTSSPAHR   TVAPSHSPFSGQEGRGPGCISPGRISGPAGR   LVLQHPTGTSPTEAKRKVPPGPPEGHPTSPVT   SPRPTPAPPRHASSGNSSVCFSKKTCRWEKK   SFVLMELAYWQDRMFF   SPRPTAPPRHASSGNSSVCFSKKTCRWEKK   SFVLMELAYWQDRMFF   AKSPLPLO*IQWRN/NGSLKLRLPGFK*FTCLG   LLSSWDYRSLPPR*PNNFCLVELGFHHYDQAG   LKLLTSSALPALASQSALTGMSHRIWPLPLLR   RPPVIRIRAPPQRLPGNLGSHSHRIWPLPLLR   RPPVIRIRAPPQRLPGNLGSSSSEGP/PQLLFVPR   RPPAPKAGAGAGASSEGSKDRGTGSLPPPSQRPLGFS   PEGAGPSPPPGGRGGSSSSEGP/PQLLFVPR   RFPAPKAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	ļ		ļ	]			HGELRRHWDRLA*GPDATEGALGASFEHEG
OPCHITGHPLGREAEPAAAGAAPALAHRPFF		1	İ	<b>i</b> 1			GQQPPADLTVQADTLHRPSARLGGAHRACPK
OPCHITGHPLGREAEPAAAGAAPALAHRPPF	l	j	J	1			RRPHRVLWRWARGAWAWRCOAREKOETOG
WPPTHPVLAGGVHFPAAG/IGGCVEVPYSVN	ł		1	i			
WPPTHPVLAGGVHFPAAG/IGGCVEVPYSVN	1						ARTGSTE\PGPCWRPIRHCRRDPLWTPTLC\RD
WMGTKSH*A VLPPPPSIGGQGLPEGWGLE   KGGLPPGIPPGILTGPWSMRPVTPSFAHIR   TVAFSHSPFSQGGRGPHGCHSPGRSGPAGR   LVLQHFTGTSPTEAKRK VPPGPPEGHFTSPVT   SPRPPTAPPRIHPASSGNSSVCFSKKTCRWEKK   SPVLMELAYWQDRMFF   AKSPLPLG*PQWRNLGSLKLRLPGFK*FTCLG   LLSWDYRSLFPRPYNFCIL VELGFHHVQAG   LKLLTSSALPALASQSAEITGMSHRIWPLPLLR   RPPVIRRAPPQRLPFNLITSLKALSPNMATF   A 3859 2 393   ALRKTRDGJRAGQPAASWKGTNNYPWR   LEMAGRPGSQEQSKDRGTGSLPPSQRPLGPS   PEGAGPSPPPGIPRGGGSSSEGPPQLLFVPR   RFPAFKKGLPSDFFHKWGLNFYHLIGGEDSQ   VPIL   VPIL   RPPVIRRAPPQRLPFNLITSLKALSPNMATF   LEMAGRPGSQEQSKDRGTGSLPPSQRPLGPS   PEGAGPSPPPGIPRGGGSSSEGPPQLLFVPR   RFPAPKKGLPSDFFHKWGLALVAQAG   EQ*RDLSSLQPPPGGRFFFLLWSLALVAQAG   EQ*RDLSSLQPPPGGRFFFLLWSLALVAQAG   EQ*RDLSSLQPPPGFF*FSCLSLPSSWDYRCP   LPCLANFVFLVETGFFHHVGQADLKLLTSGDP   PTSASSAGITGVSHRAWFRHFLYWKTFFL   APSQISVAFLYAAJDKLFEKEI*KKIPFIDASDKI   KIGNLTKEVKYLYTENYITIMKEIK/JDKW   KDILY*WIGKINI*KWRYPKAIPTKIP   MTFFTEIEKSIIKFTWNHKKPPNTQSNEQKP*S   FCSLLWVFGGFLWFLAVFNQSNEQKP*S   FCSLLWVFGGFLWFLAVFNQSNEQKP*S   FCSLLWVFGGFLWCLCARSGFGRRLCFLEP   TSNAQNVPPDLAICCFVLEQSLSVRALQEML   ANTVEAGVESSQGGHRTLJYGHAILLRHAH   SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE   ACWWTMHPASKQRSEGEKVRVGDDIILVSVS   SERYLHLSTASGELQVDASFMQTLWNMPIC   ACWTMHPASKQRSEGEKVRVGDDIILVSVS   SERYLHLSTASGELQVDASFMQTLWNMPIC   ACWTMHPASKQRSEGEKVRVGDDIILVSVS   SERYLHLSTASGELQVDASFMQTLWNMPIC   ACWTMHPASKQRSEGEKVRVGDDIILVSVS   SERYLHLSTASGELQVDASFMQTLWNMPIC   ACWTMHPASKQRSEGEKVRVGDDIILVSVS   SERYLHLSTASGELQVDASFMQTLWNMPIC   ACWTMHPASKQRSEGEKVRVGDDIILVSVS   SERYLHLSTASGELQVDASFMQTLWNMPIC   ACWTMHPASKQRSEGEKVRVGDDIILVSVS   SERYLHLSTASGELQVDASFMQTLWNMPIC   ACWTMHPASKQPLRVHVTIGQYLALIEDQG   ACWTMHPASKQPLRVHVTIGQYLALIEDQG   ACWTMATSCEKEKLDVAPKR   ACCHTMATSCEKEKLDVAPKR   AUSTMATSCETTSSKEKLDVAPKR   AUSTMATSCETTSSKEKLDVAPKR   AUSTMATSCETTSSKEKLDVAPKR   AUSTMATSCETTSSKEKLDVAPKR   AUSTMATSCETTSSKEKLDVAPKR   AUSTMATSCETTSSKEKLDVAPKR   AUSTMATSCETTSSKEKLDVAPKR   AUSTMATSCETTSSKEKLDVAPKR   AUSTMATSCETTSSKEKLDVAPKR   AUSTMATSCETTSSKEKLDVAPKR   AUSTMATSCETTSSKEKLDVAPKR   AUSTMATSCETTSSKEKLDVAPKR   AUSTMATSCETTSSKEKLDV	l .	•	Ι.				WPPTHPVLAGGVHFPAAG/IGGCVEVPVSVN
KGGGLPGIPPOGLITGPWSMRPVTPSFAHR   TVAPSHSPPSGQEGRGPHGCHSPGR\SGP\AGR     LVLQHPTGTSPTEAKRKVPPGPPEGHFTSPVT     SPRPPTAPPRHPASSGNSVCFSKKTCRWEKK     SPVLMELAY WQDRMFF     496	[	1	(	1			VMGTKSH*AVLPPPPSTGPGGOGLPEGWGLE
TVAPSISPESGQEGREPHICHSPERISGPAGR		1	ł				KGEGLPPGIPPPGLLTGPW\SMRPVTPSFAHIR
496 1846 A 3849 830 442 AKSPLPLG-IQWRNLGSLKLIRLFGFK-FTCLG LLSSWDYRSLPFRVNFCILVELGFHHVDQAG LKILITSSALPALASQSAETIGMSHRIWPLPLIR RPPVIRIRAPPQRLFFNLITSLKALSPMATF 497 1847 A 3859 2 393 ALRKITRDGIARTGAQPAASWKGTNNYPWR LEMAGRPGSQEQSKDRGTGSLPPSQRPLGFS PEGAGPSPPPPGIPRGGGSSSSEGP/PQLLFVR RPPAKKGLPSDITPHSKAPPTPHLILGGEDSQ VPIL 498 1848 A 3860 253 634 KNASTVYSSQGDFKSFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGFK*PSCLSLPSSWDNYRCP LPCLANF-*PL-VETGFHHVQADLKILTSGDP PTSASESAGITQVSHRAWPRHFLYWKTFFL 499 1849 A 3863 423 263 AFSQISVAFLYAA/DKLFEKEI*KKPFIIAS/DKI KIGINLTKEVKLYTENYTILMKEIK/DTDKW KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP MTFFTEIEKSIIKFIWNHKKPPNTQSNIEQKE*S FCSILWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNI  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGHRTILLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEML ANTVEAGVESSQGGHRTILLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEML ANTVEAGVESSQGGHRTILLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEML ANTVEAGVESSQGGHRTILLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEML CRWTMHPASKQRSEGEKVRVGDDILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFYTGGGTVLRLFHGHMDECITISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWQQPLRVRTHVTTGQYLALTEDQG LVVVDASKAHKATSFCRISKEKLDVAFKR	1	i i	ĺ	i l			TVAPSHSPFSGOEGRGPHGCHSPGR\SGP\AGR
SPRPTTAPPRHPASSGNSSVCTSKKTCRWEKK   SFVLMELAYWQDRMFF	l .	]	1	1			LVLQHPTGTSPTEAKRKVPPGPPEGHPTSPVT
SFVLMELAYWQDRMFF			1				SPRPPTAPPRHPASSGNSSVCFSKKTCRWEKK
497 1847 A 3859 2 393 ALRITRADGIARTORY 498 1848 A 3860 253 634 KNASTVYSSQGDPKSFFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGFK*PSCLSLPSSWDVYRCP LPCLANF*LVETGFHHVVQQADLKLLTSGDP PTSASESAGITGVSHRAWPRHFLYWKTFFL 499 1849 A 3863 423 263 APSQISVAFLYAA/DKLFEKEI*KKIPFIAS/DKI KIGINLTKEVKJYTENYIILMKEIK/DTDKW KULLY*WIGKINI*KMSTPPKAYPFNAIPTKIP MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNL 500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRI.CFLEP TSNAQNYPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGIRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDILLVSVS SERYLHLSTASGELQVDASFMQTIL WNMPPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR				Ł.			
LLSSWDYRSLPPRPVNFCIL VELGFHHVDQAG LKLLTSSALPALASQSAEITGMSHRIWPLPLLR RPPVIRIRAPPQRLPFINLITSLKALSPNMATF RPPVIRIRAPPQRLPFINLITSLKALSPNMATF LEMAGRPGSQCSKDGTGSLPPPSQRPLGPS PEGAGPSPPPGIPRGGGSSSSEGP/PQLLFVPR RFPAPKKGLPSDTPHSKAPPTHLILGGEDSQ VPIL  498 1848 A 3860 253 634 KNASTVYSSQGDPKSFFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGFK*FSCLSLPSSWDVYRCP LPCLANF*PLVETGFHHVGQADLKILTSGDP PTSASESAGITGVSHRAWPRHFLYWKTFPL  499 1849 A 3863 423 263 APSQISVAFLYAA/DKLFEKEI*KKIPFINASDKI KIGINLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP MTFFTEESKIRFTWNHKEPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMDFSISVKNVIGI LVGIALNL  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGIRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGGLQVDASFMQTLWMMPPIC SRCEEGFVTGGHVLRLFHGHMDECLTISSPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR	496	1846	Ā	3849	830	442	AKSPLPLG*IQWR/NLGSLKLRLPGFK*FTCLG
497 1847 A 3859 2 393 ALRKITRNDEPLLTR APPVIRIRAPPORLEPINLITSLKALSPNMATE ALRKITREDGIARTGAQPAASWKGTINNYPWR LEMAGRPGSQEQSKDRGTGSLPPPSQRPLGPS PEGAGPSPPPPGIPRGGGSSSSEGP/PQLLFVPR RFPAPKKGLPSDITPHSKAPPTPHLILGGEDSQ VPIL  498 1848 A 3860 253 634 KNASTVYSSQGDPKSFFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGPK*PSCLSLPSSWDYRCP LPCLANP*FLVETGFIHIVGQADLKLLTSGDP PTSASESAGITGVSHRAWPRIHFLYWKTFFL  499 1849 A 3863 423 263 APSQISVAFLYAADKLFEKEI*KKIPFILASDKI KIGINLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKNI*KMSTPPKATYRFNAIPTKIP MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNY* LVGIALNY*OFLOREPTPARPQPSRPARSPLPIFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKRQLKLCLAAEGFGNRLCFLEP TSNAGNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDILVSVS SERYLHSTASGELQVDASFMQTLWMMPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR	ĺ	<u> </u>	l				LLSSWDYRSLPPRPVNFCILVELGFHHVDOAG
497 1847 A 3859 2 393 ALRKTRRDGIARTGAQPAASWKGTNYPWR LEMAGRPGSQEQSKDRGTGSLPPPSQRPLGPS PEGAGPSPPPPGIPRGGGSSSSEGP/PQLLFVPR RFPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ VPIL  498 1848 A 3860 253 634 KNASTVYSSQGDPKSFFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGFK*FSCLSLPSSWDYRCP LPCLANFI*FLVETGFHHVGQADLKLLTSGDP PTSASESAGITGVSHRAWPRHFLYWKTFFL  499 1849 A 3863 423 263 APSQISVAFLYAA/DKLFKEI*KIPFILAS/DKI KIGINLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP MTFFTEIEKSIIKFIWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNL  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR		ĺ					
497 1847 A 3859 2 393 ALRKTRRDGIARTGAQPAASWKGTNNYPWR LEMAGRPGSQEQSKDRGTGSLPPPSQRPLGPS PEGAGPSSPPPPGPRGGGSSSSEGP/PQLLFVPR RFPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ VPIL  498 1848 A 3860 253 634 KNASTVYSSQGDPKSFFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGFK*FSCLSLPSSWDVYRCP LPCLANF\*FLVETGFHHVGQADLKLLTSGDP PTSASESAGITGVSHRAWPRIHFLYWKTFFL  499 1849 A 3863 423 263 APSQISVALYAAJØKLFEKEI*KKIPFILASJØKI KIGINLTKEVKYLYTENYIILMKEIKJØTØKW KDILY*WIGKINI*KMSTPPKAIYFNAIPTKIP MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE\*S FCSILLWYFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNI.  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTIDKLAFDVGLQEDATGE ACWTHHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLYYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR		1 1		1 1			RPPVIRIRAPPORLPFNLITSLKALSPNMATF
LEMAGRPGSQEQSKDRGTGSLPPPSQRPLGPS PEGAGPSPPPPGIPRGGGSSSSEGP/PQLLFVPR RPPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ VPIL  498 1848 A 3860 253 634 KNASTVYSSQGDPKSFFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGFK*FSCLSLPSSWDYRCP LPCLANF\*FLVETGFFHHVGQADLKLLTSGDP PTSASESAGITGVSHRAWPRHFLYWKTFFL  499 1849 A 3863 423 263 APSQISVAFLYAA/DKLFEKEI*KKIPFILAS/DKI KIGINLTKEVKYLYTENYITLMKEIK/DTDKW KDILY**KIGKINI*KMSTPPKAIYRFNAIPTKIP MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMN*MIDFSISVKNVIGI LVGIALNI.  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGHRTLLYGHAILRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR	497	1847	Α	3859	2	393	ALRKTRRDGIARTGAOPAASWKGTNNYPWR
PEGAGPSPPPPGIPRGGGSSSSEGP/PQLLFVPR RPPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ VPIL 498 1848 A 3860 253 634 KNASTVYSSQGDPKSFFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGFK*FSCLSLPSSWDIYRCP LPCLANF\*FLVETGFHHVQQADLKLLTSGDP PTSASESAGITGVSHRAWPRHHFLYWKTFFL 499 1849 A 3863 423 263 APSQISVAFLYAADKLFKEI*KKIPFILAS/DKI KIGINLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKINI*KMSTPPKATYRFNAIPTKIP MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALIN 500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNYPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWIMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQILWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR		<b>}</b>		1			LEMAGRPGSOEOSKDRGTGSLPPPSORPLGPS
RFPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ VPTI.  498 1848 A 3860 253 634 KNASTVYSSQGDPKSFFFLLRWSLALVAQAG E0+RDLSSLQPPPPGFK*FSCLSLPSSWDVYRCP LPCLANF*FLVETGFHHVGQADLKLLTSGDP PTSASESAGITGVSHRAWPRIHFLYWKTFFL  499 1849 A 3863 423 263 APSQISVAFLYAADKLFEKEI*KKIPFIIASDKI KIGINLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKINI*KMSTPPKATYRFNAIPTKIP MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNL  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR		]		j j	j		
498 1848 A 3860 253 634 KNASTVYSSQGDPKSFFILRWSLALVAQAG EQ*RDLSSLQPPPPGFK*FSCLSLPSSWD\YRCP LPCLANF\*FILVETGFHHVGQADLKLLTSGDP PTSASESAGTIGVSHRAWPRIHFLYWKTFFIL KIGNLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKIN\*KIGNLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKIN\*KSIPFKAIYRFNAIPTKIP MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNL  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKILCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTYEAGVESSQGGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR		!		1 1			
498 1848 A 3860 253 634 KNASTVYSSQGDPKSFFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGFK*FSCLSLPSSWD\YRCP LPCLANF\*F1.VETGFHHYGQADLKLLTSGDP PTSASESAGITGVSHRAWPRHFLYWKTFFL  499 1849 A 3863 423 263 -APSQISVAFLYAA/DKLFEKEI*KKIPFILAS/DKI KIGINLTKEVKYLYTENYIILMKEIK/DTDKW KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S FCSILLWVFQGFLWFHMNFMIDFSISVKNVIGI LVGIALNL  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHTILYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCIEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR		}	j	] ]			
EQ*RDLSSLQPPPPGFK*FSCLSLPSSWD\YRCP LPCLANF\*FLVETGFHHVGQADLKLLTSGDP PTSASESAGITGVSHRAWPRIHFLYWKTFFL  499 1849 A 3863 423 263 APSQISVAFLYAA/DKLFEKEI*KKIPFILAS/DKI KIGNLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNL  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR	498	1848	Α	3860	253	634	
LPCLANFA*FLVETGFHHVGQADLKILTSGDP PTSASESAGITGVSHRAWPRIHFLYWKTFFL  499 1849 A 3863 423 263 -APSQISVAFLYAA/DKLFEKEI*KKIPFIIAS/DKI KIGINLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNL  500 1850 A 3865 2 15246 LPRGCLWCLQRSPIPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMIDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR	,	1 ,					
499 1849 A 3863 423 263 APSQISVAFLYAA/DKLFEKEI*KKIPFILAS/DKI KIGINLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKINI*KMSTPPKAIYFFNAIPTKIP MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNI.  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTILWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQVLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR		. 1		1 1			
499 1849 A 3863 423 263 APSQISVAFLYAA/DKLFEKEI*KKIPFIIAS/DKI KIGINLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP MTFFTEIEKSIIKFIWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNI.  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMIPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTILVNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQVLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR	•			1 1	1		
KIGINLTKEVKYLYTENYITLMKEIKJÖTÖKW KDILY*WIGKINI*KMSTPPKAIYFNAIPTKIP MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNI.  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR	499	1849	A	3863	423	263	
KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP MTFFTEIEKSIIKFIWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNI.  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR							•
MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNI.  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR							
FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNI.  500 1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR			1	[ '		ľ	
LVGIALNI  1850 A 3865 2 15246 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR		·				1	
1850 A 3865 2 LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR					l	İ	
DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR	500	1850	A	3865	<del>-</del>	15246	
VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR			•-	5005	- 1		DI DDWA COI DIMODA ECEDENCEI DEDONIO
TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMIPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTILWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR				!!!	i	1	
ANTVEAGVESSQGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR					l		
SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR					l	l	
ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR					i	ļ	
SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR					1	Í	
SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR					ľ		ACW W 1 MHPASKQRSEGEK VRVGDDIILVSVS
DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR				ĺ	Į.	. [	DECEMENT ASSET OF THE PROPERTY
WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR					į	ļ	
LVVVDASKAHTKATSFCFRISKEKLDVAPKR				1	ı	1.	
					I		
DVEGMGPPEIKYGESLCFVQHVASGLWLTYA					I	į	
		L		Ll	<u>_</u> <u>l</u>		DVEGMGPPEIKYGESLCFVQHVASGLWLTYA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	·		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i				residue of	sequence	Y=Tyrosine, X=Uuknown, *=Stop codon,
Ĭ	]			peptide		/=possible nucleotide deletion, \=possible
<b> </b>				sequence		nucleotide insertion
i						APDPKALRLGVLKKKAMLHQEGHMDDALSL
1						TRCQQEESQAARMIHSTNGLYNQFIKSLDSFS
i				*		GKPRGSGPPAGTALPIEGVILSLQDLIIYFEPPS
						EDLQHEEKQSKLRSLRNRQSLFQEEGMLSMV
l i	i i					LNCIDRLNVYTTAAHFAEFAGEEAAESWKEI
<b>i</b> 1						VNLLYELLASLIRGNRSNCALFSTNLDWLVS
1						KLDRLEASSGILEVLYCVLIESPEVLNIIQENHI
1						KSIISLLDKHGRNHKVLDVLCSLCVCNGVAV RSNQDLITENLLPGRELLLQTNLINYVTSIRPN
1						IFVGRAEGTTQYSKWYFEVMVDEVTPFLTAQ
						ATHLRVGWALTEGYTPYPGAGEGWGGNGV
1						GDDLYSYGFDGLHLWTGHVARPVTSPGQHL
j						LAPEDVISCCLDLSVPSISFRINGCPVQGVFESF
	·	-		. 1		NLDGLFFPVVSFSAGVKVRFLLGGRHGEFKF
						LPPPGYAPCHEAVLPRERLHLEPIKEYRREGP
						RGPHLVGPSRCLSHTDFVPCPVDTVQIVLPPH
						LERIREKLAENIHELWALTRIEQGWTYGPVRD
1				·		DNKRLHPCLVDFHSLPEPERNYNLQMSGETL
						KTLLALGCHVGMADEKAEDNLKKTKLPKTY
	İ					MMSNGYKPAPLDLSHVRLTPAQTTLVDRLAE
1					-	NGHNVWARDRVGQGWSYSAVQDIPARRNPR
1 1						LVPYRLLDEATKRSNRDSLCQAVRTLLGYGY
						NIEPPDQEPSQVENQSRCDRVRIFRAEKSYTV
						QSGRWYFEFEAVITGEMRVGWARPELRPDV
1 1						ELGADELAYVFNGHRGQRWHLGSEPFGRPW
				i		QPGDVVGCMIDLTENTIIFTLNGEVLMSDSGS
					i	ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQD VSSLRFFAICGLQEGFEPFAINMQRPVTTWFS
j l		j				KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCLR
						LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFR
				ļ		CTAGATPLAPPGLQPPAEDEARAAEPDPDYE
1	0.0	1				NLRRSAGGWSEAENGKEGTAKEGAPGGTPQ
1 1			i			AGGEAQPARAENEKDATTEKNKKRGFLFKA
1 . 1				•		KKVAMMTQPPATPTLPRLPHDVVPADNRDD
						PEIILNTTTYYYSVRVFAGQEPSCVWAGWVT
1		ļ			]	PDYHQHDMSFDLSKVRVVTVTMGDEQGNV
			1			HSSLKCSNCYMVWGGDFVSPGQQGRISHIDL
j		l	1			VIGCLVDLATGLMTFTANGKESNTFFQVEPN
		1	1		]	TKLFPAVFVLPTHQNVIQFELGKQKNIMPLSA
		- !	}	j	ł	AMFQSERKNPAPQCPPRLEMQMLMPVSWSR
		i	1	I	1	MPNHFLQVETRRAGERLGWAVQCQEPLTMM
		l				ALHIPEENRCMDILELSERLDLQRFHSHTLRL
				į	. [	YRAVCALGNNRVAHALCSHVDQAQLLHALE
]		J	1		. 1	DAHLPGPLRAGYYDLLISIHLESACRSRRSML
	·		ĺ			SEYIVPLTPETRAITLFPPGRSTENGHPRHGLP
1		ł		×	•	GVGVTTSLRPPHHFSPPCFVAALPAAGAAEAP
		l				ARLSPAIPLEALRDKALRMLGEAVRDGGQHA
1 1		ł		i	ł	RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE
			1		ł	DVKQILKMIEPEVFTEEEEEEDEEEEGEERDEE
[ ]		1	. [		}	EKEEDEETAQEKEDEEKEEEBAAEGEKEEG LEEGLLQMKLPESVKLOMCHLLEYFCDOELO
1	j	J	J	1		HRVESLAAFAERYVDKLOANORSRYGLLIKA
	1			1	I	FSMTAAETARRTREFRSPPQEQINMLLQFKDG
	1			ļ	ì	TDEEDCPLPEEIRQDLLDFHQDLLAHCGIOLD
			i	į	ł	GEEEPEETTLGSRLMSLLEKVRLVKKKEEK
[ ]	. [	1	- 1	ľ	ì	PEEERSAEESKPRSLQELVSHMVVRWAQEDF
1	- 1				l	VQSPELVRAMFSLLHRQYDGLGELLRALPRA
	ļ	- 1	ļ	İ	l	YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQE
1	]	j	}	l	J	ENLMIQSIGNIMNNKVFYQHPNLMRALGMHE
	[				ļ	TVMEVMVNVLGGGESKEIRFPKMVTSCCRFL

NO: of mucle ootde of the period of the peri	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
soq- uence  914  914  916  1916  1916  1917  1917  1917  1918  1918  1918  1918  1919  191	NO: of	NO: of	hod		beginning	nucleotide	
uence    9/1496   gr biff in gr b first   gr biff in gr biff   gr	nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
September   Sept	eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
September   Sept	seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
residue of peptide sequence    y	nence			914	ng to first	acid residue	
residue of peptide sequence    Poptide in uncleotide deletion, "possible nucleotide deletion, "possible nucleotide deletion, "possible nucleotide insertion    CYFCRISRONGRSMPHLSYLLENSGIGLGIM   QGSTPLDVAAASVIDNNELALALQEQDLEKV   VSYLAGGCJQSCPHLVAKGYPDIGWFCPG   ERYLDFLRFAVFVNGESVEENANVVVKLIJE   RFECFGPALREGEGSGILAAITERARISEDPAR   DOPGRRORRREHFGEEPPEENRVHLGHAIMS   FYAALDILGRCAPEMHLIQAGKGEALRRAI   IRSLVPLEDLVGISLPLOJPTIGKDGALVOPK   MASS-YVDIKASMVLT,DRAYVERDOPDILL   VLDVGFLPDMRAAASLDTATFSTTEMALAV   NRYLCLAVPLITIKAPLFARTTEHRAIMVDS   MAHTVYRLSKGRSLTKAQRUVIEDCLMSLCR   TRENGLGELLRRAI VOTELYRMAMCLC.   ARAGALPPTYVDASYSSKARKATVDAEGNT   DRAWNSVEEDLSTRIKAPROVIEDLAMSCC.   ARAGALPPTYVDASYSSKARKATVDAEGNT   DRAWNSVEEDLSTRIKAPROVIEDLAMSCC.   ARAGALPPTYVDTITAFRER RANGE   KOKENYRWPRICSI KAAILAWBYTEK AREGE   EKTEKKKTAKUSOS ONTVDREGVPROPPDL   SAVTLSRELQAMAEGLAENYHNTWGRKKKQ   ELEAKGGGTIPPLLYVTDITTAREK RANGE   EEKTEKKKTAKUSOS ONTVDREGVPROPPDL   SAVTLSRELQAMAEGLAENYHNTWGRKKKQ   ELEAKGGGTIPPLLYVTDITTAREK RANGE   EKSPHEGGLEFFAKILLPLINQYFTNHCLYFLS   ERSPHEGGEFFAKILLPLINQYFTNHCLYFLS   FAFOFIOQULE WINDISCEPIAHLEAVSSKRV   ERSPHEGGEFFAKILLPLINQYFTNHCLYFLS   ERSPHEGGEFFAKILLPLINQYFTNHCLYFLS   FAFOFIOQULE WINDISCEPIAHLEAVSSKRV   EKSPHEGGEFFAKILLPLINQYFTNHCLYFLS   ENTAKLYGGGHASNKEREMITSLFCKLAALV   RIRVSLFGTDAPAVVNCLHILARSLDARTVM   SKGPETVKAGGRAFFSKEEMSTENKENSTRIKLG   KVSQARTQVKGVQNLTYTTVALLPVLTTLF   QHAQQFGDDVLLDDVOSCYRTLCSTYSLG   TTKNTYYEELRPALGECLARLAAAMPVAFLE   POLNEYNAGSVYTKSKERRAILGFNSVERM   CPDIPVLERLMADIGGLAESGARYTEMPHVE   ENTAKLGGGGGARASKERREMTSLFCRIKK   AGKVVSBERGLALBAKAEAQHGELLVEDIFFS   VLCDILYALTYLLWFLYNDRAQWITTERNS   AEELFMWVGEITYWSKSHPKREGSNFVVQ   NENNMSFLTADMSKSKAMAAGGISGOOD   RCAATGOOLTI ACTTRAV TOTREGREKE   AGKVVSBERGLALBAKAEAQHGELLVEDIFFS   AEELFMWVGEITYWSKSHPKREGONFVV   NENNMSFLTADMSKSKAMAAGGISGOOD   RCAATGOOLTI ACTTRAV TOTREGREKE   ALTEKSKLDEDYLVMAYADDFTOLITERPS   AEELFMWVGEITYWSKSHAWALAGDERGELLE   DADDPEKUTRINQESVAAAWULTEDISSPERMIDD   ISAGGGGGEFEEVEVSFERKOME CONTRINGKSKAM   SVARQVYNSLTETQIGTTTTNIIGTTVDYLL   EGHNNDPONYLKTQOFTCHO					amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
sequence	1	Į	}		residue of	sequence	
CYFCRISKONORSMFDHISYLLENSGIGLGM QSTPLDVAAASVUDNBELALALQEQDLEKV VSYLAGGGLQSCPMLVAKGYPDIGWKPCGG RYLDFLRFAVPVNORSSYEENANVVWLLIR KFECFGPALRGGGSGLLAAIEARISEDPAR DCPGREDERREHEGEPFERREVHLIGHAMS FYAALDLLGRCAFEMHLIQAGKGEALRIRAI LESLVPLEDLVGGIELPLOFTIGKOGALVOPK MSASFYDEHASMVLFLDRYVGENODFILH VLDVGFLPDMRAAASLDTATTSTTEMALAV NEYLCLAAVPLEDLVGGLAVOPK MSASFYDEHASMVLFLDRYVGENODFILH VLDVGFLPDMRAAASLDTATTSTTEMALAV NEYLCLAAVPLEDLVGGLAVGENOFELH TRIKLFWGIPDSLAHKKYDFELJAMSLCR VRPSMLQHELRRLYFDVPLNEPAKMPLKLL THINTERCWKYYCLFTWAADPROVERSELH TRIKLFWGIPDSLAHKKYDFELJRMAAMPLC AAGOPPOYNDASYSKAEKKANPLKLL THINTERCWKYYCLFTWAADPROVERSELH TRIKLFWGIPDSLAHKKYDFELJRMAAMPLC AAGOPPOYNDASYSKAEKKANPLKAL TRIKLFWGIPDSLAHKKYDFELJRMAAMPLC AAGOPPOYNDASYSKAEKKANPLKAR BKLQNNWSYGERIDELLKTHPAHERVAFE BKLQNNWSYGERIDELLKTHPAHERVAFE BKLQNNWSYGERIDELLKTHPAHERVAFE BKLGENTRYWFIRESLKAMMAWEWITERVKTES EKKTEKKYTAKISGSAQTYDPREGYMPOPPOL SAYLERELQAMARQLABYHTHYWGRKKQ ELRAKGGTHPLJAYFTTTAKEKANDREKA QELLKFLQNNWYAYTRCLABWHTHYWGRKKQ ELRAKGGTHPLJAYFTTTAKEKANDREKA QELKFLQNNWYAYTRCLABWHTHYWGRKKQ EKSHELQOKATYACHAGANAYAYAGE EKSHELQOKATYACHAGANAYAYAGE EKSHELQOKATYACHAGANAYAYAGE BKSHELGOKATYACHAGANAYAYAGE BKSHELGOKATYACHAGANAYAYAGE BKSHELGOKATYACHAGANAYAYAGE BKSHELGOKATYACHAGANAYAYAGE BKSHELGOKATYACHAGANAYAYAGE BKSHELGOKATYACHAGANAYAGE BKSHELGOKATYACHAGANAYAYAGE BKSHELGOKATYACHAGANAYA	1	1			peptide	1	/=possible nucleotide deletion, \=possible
QGSTFLDVAAASVIDNNELALAQEQDLEV VSYLAGCGLQSCEMLVAKGYPDIGWKPCGG ERYLDFLRFAVFVNGESVEENANVVVRLIL KPCCFGALREGGSGLLAATEARISEDPAR DGFGIRRDRREHIGGEFFFENRVHLGHAMS FYAALIDLLGRCAPEMHLIQAGALRIRAI LRSLVPLEDLVGIISLPLQIFTLGKDGALVGFW MASSFVPDHASANVLFLDRVYGIENQDFLIH VLDVGFLPDMRAAASLDTATFSTTEMALAV NRVLCLAVIL-HITCAPLFAGTEMAMVDS MLHITVYRLSRGRSLTKAQEDVIEDCLMSLCR YRPSMAQHLLRRL VPDVPLLNERAMVDS MLHITVYRLSRGRSLTKAQEDVIEDCLMSLCR YRPSMAQHLLRRL VPDVPLLNERAMVELL TNHYERCWKYYCLFTGWANFGVTSEELHL TRILFWGFDSLAHKKYPDELLYMAMPLLL TRILFWGFDSLAHKKYPDELLYMAMPLL TRILFWGFDSLAHKKYPDELLYMAMPLL AGAGALPPDVVDASYSSKAEKKATVDAEGM DFRPVETLNVIIPEKLDSFINKFAKHPLSL TRILFWGFTSSLAKKYPDELYMAMPLL AGAGALPPTVVDASYSSKAEKKATVDAEGM DFRPVETLNVIIPEKLDSFINKFAKHPLSL TRILFWGFTSLAKKYPDELYMAMPLL TRILFWGFTSLAKKYPDHATEACHYVV NEINMSFLTADNKSKMAKAGDIGGGGSGG TKKKRAGDRYSVGTSLINGLINGNINGLDASWM TRILFWGFTSLAKKYPDHATEACHYVV NEINMSFLTADNKSKMAKAGDIGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	ì	•			sequence		nucleotide insertion
QGSTFLDVAAASVIDNNELALAQEQDLEV VSYLAGCGLQSCEMLVAKGYPDIGWKPCGG ERYLDFLRFAVFVNGESVEENANVVVRLIL KPCCFGALREGGSGLLAATEARISEDPAR DGFGIRRDRREHIGGEFFFENRVHLGHAMS FYAALIDLLGRCAPEMHLIQAGALRIRAI LRSLVPLEDLVGIISLPLQIFTLGKDGALVGFW MASSFVPDHASANVLFLDRVYGIENQDFLIH VLDVGFLPDMRAAASLDTATFSTTEMALAV NRVLCLAVIL-HITCAPLFAGTEMAMVDS MLHITVYRLSRGRSLTKAQEDVIEDCLMSLCR YRPSMAQHLLRRL VPDVPLLNERAMVDS MLHITVYRLSRGRSLTKAQEDVIEDCLMSLCR YRPSMAQHLLRRL VPDVPLLNERAMVELL TNHYERCWKYYCLFTGWANFGVTSEELHL TRILFWGFDSLAHKKYPDELLYMAMPLLL TRILFWGFDSLAHKKYPDELLYMAMPLL TRILFWGFDSLAHKKYPDELLYMAMPLL AGAGALPPDVVDASYSSKAEKKATVDAEGM DFRPVETLNVIIPEKLDSFINKFAKHPLSL TRILFWGFTSSLAKKYPDELYMAMPLL AGAGALPPTVVDASYSSKAEKKATVDAEGM DFRPVETLNVIIPEKLDSFINKFAKHPLSL TRILFWGFTSLAKKYPDELYMAMPLL TRILFWGFTSLAKKYPDHATEACHYVV NEINMSFLTADNKSKMAKAGDIGGGGSGG TKKKRAGDRYSVGTSLINGLINGNINGLDASWM TRILFWGFTSLAKKYPDHATEACHYVV NEINMSFLTADNKSKMAKAGDIGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG					-		CYFCRISRONORSMFDHLSYLLENSGIGLGM
VSYLAGCGIQSCPMLVAKGYPDIGWIKTOGG ERYLDFLRFAVFVOGSVEENANVVRLILE KPECGRALRGEGGSGILAATEEARISEDRA DÖGGIRDRREIF GEFFFEENSVHIGHAMS FYAALIDLLGRCAPEMHLIGAGKGEALRIKAL LSL.VPLEDLVGGISDLPLQIPTIGAGKGGALRIKAL LSL.VPLEDLVGGISDLPLQIPTIGAGKGGALVQFK MSASFVPDHKASMYLFLDRVYGIENQDFLLH VLDVGFLPDMRAAASLDATAFFRALAV NRYLCLAVIPLITKCAPLFAGTEHRAIMVDS MLHTVYRLSRGGSLTAQADUDECLMSLCR YRFSMI,OHLLRRLVFDVPILAGEDCLMSLCR YRFSMI,OHLLRRLVFDVPILAGEARAPLXLL THYPECWKYYLLFTGWANPICYEBELHL TRKLFWGFFDSLAHKKYDPELYRMAMPCLCL AIAGALFPDVVDASYSSKAEKATVDAEGEN DEPVETI NVIIPEKLDSFINKFAEVTHEKWAF DKIQNINWSYGENDIDELLTHEIP KYFTISEL KDEETYRWPIKESLKAMMAWWTEKARFER EKTEKKKTAKSSYSSKAEKATVDAEGEN DKONNWSYGENDIDELLTHEIP KYFTISEL KDEETYRWPIKESLKAMMAWWTEKARFER EKTEKKKTAKSGSAOTYDPREVARPOPPDL SAVTLSRELQAMAEQLENYHNTWGRKKKQ ELAKGGGTIFFLVYPDITLAKRANPEQPDL SAVTLSRELQAMAEQLENYHNTWGRKKKQ ELAKGGGGIFFLVYPDITLAKRANPEQPDL SAVTLSRELQAMAEQLENYHNTWGRKKKQ ELAKGGGIFFLVYPDITLAKRANPEQSERV EKSPHEORIKFFAKILLPLINQYFTHHCLYSLS TRAVVLGGGGIBANKKEKEMITSPLCKLAALV RIREVSLGYGTDAPAVVNCLHILARSLDARTVYNC KSCPPTVKAGLASPFESASEDIEKHANPECKLAALV RIREVSLGYGTDAPAVVNCLHILARSLDARTVILL STAVLVGGGGGIBANKKEKEMITSPLCKLAALV KNOQARTOVKGYGGONLTYTTVALLPVLTLE QHAOROGODVITTTVSPERALGLENYHNURL KNOQARTOVKGYGGONLTYTTVALLPVLTLFL QHAOROGODVITTTSPSPERALGLENAAMPVAFEL POLNSYNACSSIVTTSSPPERALGLEPROPPE TLTMTTYYSEKLRPALGECLARLAAAMPVAFEL POLNSYNACSSIVTTSSPPERALGLEPROPPE TLTMTTYYSEKLRPALGECLARLAAAMPVAFEL POLNSYNACSSIVTTSSPPERALGLIPROPPE TLTMTTYTYSEKLRPALGECLARLAAAMPVAFEL POLNSYNACSSIVTTSSPPERALGLIPROPPE TLTMTTYTYSEKLRPALGECLARLAAAMPVAFEL POLNSYNACSSIVTTSSPPERALGLIPROPPE TLTMTTYTYSEKRRPATERQOPPVVQ NEINNMSTLTANNASKMAKAGDIORGGGSDG RYKKKRRODRYSVGSRLWGMALJTROPPGREB DADDPHAUERLMADIOGLAESGARYTEMPRIPTIGLERG ROEGEREEVSEKRPOPLAYON NEINNMSTLTANNASKMAKAGDIORGGGBGG RYKKKRRODRYSVGSLLWGMALJTROPPGREB DADDPHAURHLSSYRLWGMALJTROPPGREB DADDPHAURHLSSYRLWGMALJTROPPGREB DADDPHAURHLSSYRLWGMALJTROPPGREB DADDPHAURHLSSYRLWGMALDEFTOJH, VLHFSST ALTEKSKLDEDDYLYMAYADIMAKSCHLEGG GENGEREEVSENSPLWAMALDEFTOJH, VLHFSST ALTEKSKLDEDDYLYMAYADIMAKSCHLUFSSER LEGGGREBESEVSTEREKOMEKGRULVQ	1	1					
ERYLDFLRFATFVNGESVERANVVVRLIB KPECEGPALREGEGSGLLAAREARISEDPAR DOPGIRRDRREHIGEPPERNEVHLGHARMS PAALIDL.GGCSGSGLLAAREARISEDPAR DOPGIRRDRREHIGEPPERNEVHLGHARMS PAALIDL.GGCSPPMHLIGAGALRIAI LISSLYPLEDL.VGIISLPLOIPTI.GEDGALVOPLK WASSYPOPHAKSANVLFLDREVYGENODFLIK VLDVGFLPDMRAASI.DTATFSTTEMALAY NRVLCLAVLPLIKCAPLFAGTEMALWOD MLHTVYRLSRGRELTRAQEDVIEDCLMSI.CR YYEPSMLQHLLRILVFDVPLNEPAKMPLKLL TRKLFWGIPDSLAHKKYDPELYRMAMPCLC. ALGALAPPD VVDASYSKARKKATVDABGINP DREVETLINVIDEKLDSFIRKFAFYTHERWAF DIGONNWSYGENIDEELLKTHMLRFYKTESE KOKEIYRWPIKESI.KAMMWEWTIEKAREGE EKKTEKKKTAKISSORAOTYPDREGYMPQPPDL SAVTLSRELQAMAEQLAENYHNTWGRKKKK QELLKFLQMNGYAVTIGGLEMYHNTWGRKKKK QELLKFLQMNGYAVTIGGLEMYHNTWGRKKKK QELLKFLQMNGYAVTIGGLEMHELAVYSSGRV EKSPHEQEIKFFAKILLPLINQYFTHHALFAVISGRV EKSPHEQEIKFFAKILLPLINQYFTHHALFAVISGRV EKSPHEQEIKFFAKILLPLINQYFTHHALFAVISGRV EKSPHEQEIKFFAKILLPLINQYFTHHALFAVISGRV EKSPHEQEIKFFAKILLPLINQYFTHHALFAVISGRV KNOQARTOVKGVGONLTYTTVALLPVLITLF QHAQOFGODVITTISSPLEGALALAV KNOQARTOVKGVGONLTYTTVALLPVLITLF QHAQOFGODVITTISSPLEGALAGANAPWAFE CDIPVLEELMADIGGLASEGARYTEMPPUT TIFMCSSYLFRWERGFEAPPSALPAGAPPP CTAYTSPHLINSLLGRILRIVINNLGIDEASWA KRIANFAQPTVSRAPPHLQSECLARLAAAMPVAFE CDPPVLEELMADIGGLASEGARYTEMPPTIGRIKR AGKVVSEEGQLALHAKARAQEGGELLVRIDERS VLCRDLYALPFILLRYVDNINGARAPAPPT CTAYTSPHINSLLGRILRIVINNLGIDEASWA KRIANFAQPTVSRAPPHLQSITTIGRIKR AGKVVSEEGQLALHAKARAQEGGELLVRIDERS VLCRDLYALPFILLRYVDNINGANGAVITEPPS AEELFRAWGERTYWSKISHPKREEONINVO NEINNMSFLTADNKSKMAKAGIGGGGGSOO RYKKKRRGGRYSVOTSLIVATLKKMLPIGLIN MCAPPTOPOLITIALKKRAPPELLOSLITARLIKHLILGURINGRUNGARAVACPRITPLYMIL TRRACOMHLESYRAAWULTEPPSPRALD LIKAGEGREEPEVSEKKPOPLHQUVLHEPST ALTERSKLDEDYLYMAXAAVACPRITPLYMIL TRRACOMHLESYRAAWULTEPSPERMID LIKAGEGREEPEVSEKKPOPLHQUVLHEPST ALTERSKLDEDYLYMAXAAVACPRITPLYMIL TRRACOMHLESYRAAWULTEPSPERMID LIKAGEGREEPERVEKKROPLINGULTRICLULL GGINNDRGNILKTYQMICTICTITTITITICTTVDYLL LIKAGABRAVLQMISACKGETGAMYSSILLKE LIDLOKMVVMILSLLEGNVNIGMARGAM SVAKCYPINSLTEYTIGGCTGNQGSLAHSEML DAVVGFLIFVAHMMKKLAQOSGGLLKEL LIDLOKMVMULSLLEGNVNIGMARGAM	į.	1					
KPECFGFALRGEGGSGLLAAIEEARRISEDRA DÖFGRRDRREIFGEFFFENVHGHAMS FYAALIDLLGRCAPEMHLIGAGKGGALRIAL LSLYPLEDLVGGISPLQOFTLGGGALVOPK MSASFVPDHKASMYLFLDEVYGEROODEL WEGELDWAG MSASFVPDHKASMYLFLDEVYGEROODEL WEGELDWAG MSASFVPDHKASMYLFLDEVYGEROODEL WEGELDWAG MSASFVPDHKASMYLFLDEVYGEROODEL WEGELDWAG MITTYRLSGRSLTKAGEDVIEDCLMSLCR TURNSMIGHLERU, FYDVPURLBERALAV NRYLCLAVLPLITKCAPLFAGTEHRAMVDS MLHTYRLSGRSLTKAGEDVIEDCLMSLCR TURNFYBRCWKYYCLFTGWANFGVTSEBELHL TINHYBRCWKYYCLFTGWANFGVTSEBELHL TINHYBRCWKYYCLFTGWANFGVTSEBELHL TINHYBRCWKYYCLFTGWANFGVTSEBELHL TINHYBRCWKYYCLFTGWANFGCYTSEBELHL TINHYBRCWKYYCLFTGWANFGCYTSEBELHL TINHYBRCWKYYCLFTGWANFGCYTSEBELHL TINHYBRCWKYYCLFTGWANFGCYTSEBELHL TINHYBRCWKYYCLFTGWANFGCYTSEBELHL TINHYBRCWKYYCLFTGWANFGCYTSEBELHL TINHYBRCWKYYCLFTGWANFGCYTSEBELHL TINHYBRCWKYYCLFTGWANFGCYTSEBELH SAVTLSBELQAMAGQLAENYHTWGRKKQ ELEAKGGTTHFLLVPYDTITAKKAADRKKQ ELEAKGGGTHFLLVPYDTITAKKAADRKKQ ELEAKGGGTHFLLVPYDTITAKKAADRKKQ ELEAKGGGTHFLLVPYDTITAKKAADRKKQ ELEAKGGGGASNIKEKEMTSLFCKLAALV RIRVSLGFTDAPAVYNCLHILASDARTVM KSGPEIVKAGLRSFPESASEDIEKMYENLRLG RIRVSLGFTDAPAVYNCLHILASDARTVM KSGPEIVKAGLRSFPESASEDIEKMYENLRLG WEGELFGWAGGONLYTTVLYUTLT QHLAOHOPGDDVLLDDVOVSCYRTLCSIYSLG TINHYYPKLRPALGGCLARLAGCAGLAPVYTTL QHLAOHOPGDDVLLDDVOVSCYRTLCSIYSLG TINHYYPKLRPALGGCLARLAGCAGLAPVYTTL TIPHLCSYLPRWERGPAPPSALPAGAPPY CTAYTSDHLNSLLGHILRIVNNLGDGASW KRLAVPGPTYRKARPELLGSTEIPTIGRIKK AGKVVSEEGQLALEKARAQEGGLLVROEPS VLCRDYALPFLLIKTYVDNNKAROWLTEPNS VLCRDYALPFLLIKTYVDNNKAROWLTEPNS VLCRDYALPFLIKTYVDNNKAROWLTEPNS VLCRDYALPFLIKTYVDNNKAROWLTEPNS VLCRDYALPFLIKTYVDNNKAROWLTEPNS VLCRDYALPFLIKTYVDNNKAROWLTEPNS VLCRDYALPFLIKTYVDNNKAROWLTEPNS VLCRDYALPFLIKTYVDNNKAROWLTEPNS VLCRDYALPFLIKTYVDNNKAROWLTEPNS VLCRDYALPFLIKTYVDNNKAROWLTEPNS VLCRDYALPFLIKTYVDNNKAROWLTEPNS VLCRDYALPFLIKTYTTINHTURCTVDTL TIRACAMPLSSERPERABLUSHEFORMIDD LSKAGEGEBEEVEVEKKPPPLHQLVLTTLP TIRACAMPLSSERPERABLUSHURGERGRANVSTILL HTGGAABWYLOMBACKGETGROWNSTILL HTGGAABWYLOMBACKGETGROWNSTILL HTGGAABWYLOMBACKGETGROWNSTILL HTGGAABWYLOMBACKGETGROWNSTILL LGESIDPYWYTSKONDAROWNSTILL LGLESSDYWWHALSLLEDNVNGMAROWY DLVYGESSN		}		ļ		ļ	1
DÖGGIRDDRREHIGGEPFERNVHLGHAMS FYAALIDLIGGCALPEMHLIQAGGGALRIRAI LRSLYPLEDLVGISLPLQIPTI.GKDGALVQFK MASRYPPHKASMVLFLDRVYGISNQDFLJK VLDVOGLPDMRAAASLDTATFSTTEMALAV NRVLCLAVLPLITKCAPLAGTEMAMVDS MLHTVYRLSGGSLTKAQEDVIEDCLMSI.CL YRPSMJ.OHLERULYDDVPLLBEMAMPLI.CL TRIFWGIPTSLAHKKYDPEL PRAMAMPLI.CL TRKLFWGIPTSLAHKKYDPEL STRAMAPLI.CL TRKLFWGIPTSLAHKKYDPEL STRAMAPLI.CL ATAGALPPDYVDASYSSKAEKKATVDAEGNP DRPVEILNVIDEKLDISTINKFAE'THER WA DKIQNINWSYGGINDELEKTEPMLPYKTFSSE KOKENTRWPIKESLKAMMAWWTIEKAREGE BEKTEKKTAKIGSGAQTYPDREMPOPPDL SAVTILSRELQAMAEQLAENYIHNTWGRKKKC ELEAKGGGTHFLLVYDTITAKEKADREKK QELLKFLQMNGYAVTRGLEMBLEAVYSSGRV EESPHEQEIKFFAKILLPLINQYFTIHCLYTH, SKOPETYKAGLISFFESSASDIEKMYENLRUL KYSQARTQVKGVGONLTYTTVALLPVLTTLP QHAQOGGDDVITTSYNLSPEALALGLPKLAALV RIRVSLIGTDAPAVVNCLHILARSLDARTVY KSOPETYKAGLISFFESSASDIEKMYENLRUL KYSQARTQVKGVGONLTYTTVALLPVLTLP QHAQOGGDDVITTSSPEALAGLEPKLESSLSGL TTKNTYVEKLRPALGECLARLAAAMPVARL PQHAQOGGDDVITTSSPEALAGAPPP CTAYTSHLINSLLGNILRIVINNLGIDEASWW KRLAYPAQOPVSRARPELLQSHIPTIGRIRKR AGKVVSEEGQLALBKARAQEGGELLVRDEFS VLCRDLYALFYLIKSYBERGELAPSPALPGAOPP CTAYTSHLINSLLGNILRIVINNLGIDEASWA KRLAYPAQOPVSRARPELLQSHIPTIGRIRKR AGKVVSEEGQLALBKARAQEGGELLVRDEFS VLCRDLYALFYLLIRYVDNINAGONGGGSGGSOO ERTKKRRGGRYSVGTSLIVATLKKMLPIGLIN MCAPTOPOLITLRIYDNINGKSKMAKAGGOGGGSOO ERTKKRRGGRYSVGTSLIVATLKKMLPIGLIN MCAPTOPOLITLRIYDNINGKSKMAKAGGOGGGSOO ERTKKRRGGRYSVGTSLIVATLKKMLPIGLIN MCAPTOPOLITLRIYDNINGKSKMAKAGGOGGGSOO ERTKKRRGGRYSVGTSLIVATLKKMLPIGLIN HTRGAABAVLQMISACKGCTGAMYSSTILL LIGGSIDAPSVRICKSVGRAVAVACPRITTLYTH TRRACOMHLESSYRAWALTRYALKOTICHTITICITOTYLL LIGGSIDAPSVRICKSVGRAVALTGRYGFRE ALTERSKLDEDVIYJAMAADDHACHERET ALTERSKLDEDVIYJAMAADDHACKERGE GENGBAEEEVEVSFEKQMEKGRLLYQQARL HTRGAABAVLQMISACKGCTGAMYSSTILL LIGGSIDAPSVRICKSVGRAVACPRITTLYTHL LIGGGRAGAEEVEVSFEKQMEKGRLLYQQARL HTRGAABAVLQMISACKGCTGAMYSSTILL LIGGSIDAPSVRICKSRAWALCPRITTLINICTVOYLL LIGGSIDAPSVRICKSRAWALCPRITTLINICTVOYLL LIGGSIDAPSVRICKSRAWALCPRITTLINICTVOYLL LIGGSIDAPSVRICKSRAWALCPRITTLINICTVOYLL LIGGSIDAPSVRICKSRAWALCPRITTLILLL LIGGSIDAPSVRICKSRAWALCPRITTLILLL LIGGSIDAPSVRICKSRAWALCP						1	
LRSLYPLEDLYGISLPLQFTICKEDGALYOPK MSASFYPDHKASMVLFLDRYGIENQDFLIH VLDVGFLPDMRAAASLDTATFSTTEMALAV NRYLCLAVLPLTTRCAPLFAGTEHRAMAVDS MLHTVYRLSRGRSLTKAORDVIEDCLMSLCR YRPSMLOHLJRRLYPDVPLTBEKAMPLKLL TNEYERCWKYYCLPTGWANFOVTSEEELHL TRKLFWGIFDSLAHKKYDFELYRMAMFCLC AIAGALPPDYVDASYSSKAKATVDAEGNF DPRPVETLNVILPEKLDSFINKFAEYTHEEWAP DKIQNNWSYGENIDELKITHPMLRPYKTFSE KDKEIYRWPKESLKAMIAWEWTIEKAKTOPEL SAVTLSRELQAMAEQLAENYHINTWORKKQ EERKGGGTHPLLYPYDTLTAKEKARDERKA QELLKRIQOMOTAYTIGLKDMELDSSISEKR FAGGLALRYLOWNGYAYTIGLKDMELDSSISEKR FAGFILQOLLEWMDISQEFLAVUNCHILLAWSGROW EKSPHEQERFFAKILLPLNQYFTINHLEAVYSGROW EKSPHEQERFFAKILLPLNQYFTINHLEAVYSGROW EKSPHEQERFAKILLPLNQYFTINHLEAVYSGROW EKSPHEQERFFAKILLPLNQYFTINHLAG KYSQARTQVKGVGQNLTYTYALLPVLTLF QHAQBIGGRODVILDDVQVSCYRILLSINSLG TIKNTYVEKLRPALGECLARLAAMPVAFLE PQLNEYNACSVYTIKSPRERAILGERYNSVERM CPDIPVLERLMADIGGLAESAGTEMPHYEE TILPMLCSYLFRWANGEGPEAPPSALPAGAPPP CTAYTSDHLINSLLONILRIVNILGIDEASWM KRILAYFAQFIVSRAFRELIGLSHFPTIGRIKKR AGKVSEEBQLAIEAKAARGGELLVRDEFS VLCRDLYALYPLLRYVDNNRAQWLTEPPS ERIFRMYGEGITYWSKSINFACGNILRIKA AGKVSEEBQLAIEAKAARGGGGLLVRDEFS VLCRDLYALYPLLRYVDNNRAQWLTEPPS ERIFRMYGEGITYWSKSINFEQNIVVQ NEINNMSFLTADNRSKMAKAGDIQSGGSDQE RTKKKRRGROFSYOYSLLVKKMILFIGDHIPTIGNILRIK KRAVPHELBYNGROFSYNGYSLTAKKMILFIGNILRIVNILF THRACMMFLESFYKAAWILTISHFEGNIPTVO NEINNMSFLTADNRSKMAKAGDIQSGGSDQE RTKKKRRGROFSYNGYSLTAKKMILFIGNILFROFWARTHINH LKHLQGKVEGSPSLWQMALYRCVFGRE DADDPEKIVRRVQEVSAVLYYLQTGHEPYKS KRAVWHELISKQRRRAVVACPRIMTINHL THRACMMFLESFYKAAWILTISHFEGNIPTVO LSKAGBQEEEBEEVEEKKPDPLHQLVLHFBRT ALTERSKLLDEDYLVMAYADIMAKSCHLEEG GENGBAEEEVVSFERKQMEKQLLLYCHFSRT ALTERSKLLDEDYLVMAYADIMAKSCHLEEG GENGBAEEEVVSFERKQMEKQLLLYCHFSRT ALTERSKLLDEDYLVMAYADIMAKSCHLEEG GENGBAEEEVVSFERKQMEKQLLLYCHFSRT ALTERSKLLDEDYLVMAYADIMAKSCHLEEG GENGBAEEEVVSFERKQMEKQLLLYCHFSRT ALTERSKLLEDGYLVMAYADDEFTQDLFRTPLQLLC EGHNNDFQNYLRTQTONTTTINICITVDYLL RQESISDFYWYYSGROWLEGGGRANWSSTIKL DAVVGFEHYYANGKAMACAGDIGGOGGBOLELKEL LDLQKDMVYNGLILLEGNVYNGMARGMY DMVSSSNYVERMILKFTOMFTATHUNGTOSTAFL DAVVGFEHYYANGKMILFTOMOTATHUNGDSQELLKEL LDLQKDMVYNGLILLEGNVYNGMARGMY		İ				İ	
LRSLYPLEDLYGISLPLQFTICKEDGALYOPK MSASFYPDHKASMVLFLDRYGIENQDFLIH VLDVGFLPDMRAAASLDTATFSTTEMALAV NRYLCLAVLPLTTRCAPLFAGTEHRAMAVDS MLHTVYRLSRGRSLTKAORDVIEDCLMSLCR YRPSMLOHLJRRLYPDVPLTBEKAMPLKLL TNEYERCWKYYCLPTGWANFOVTSEEELHL TRKLFWGIFDSLAHKKYDFELYRMAMFCLC AIAGALPPDYVDASYSSKAKATVDAEGNF DPRPVETLNVILPEKLDSFINKFAEYTHEEWAP DKIQNNWSYGENIDELKITHPMLRPYKTFSE KDKEIYRWPKESLKAMIAWEWTIEKAKTOPEL SAVTLSRELQAMAEQLAENYHINTWORKKQ EERKGGGTHPLLYPYDTLTAKEKARDERKA QELLKRIQOMOTAYTIGLKDMELDSSISEKR FAGGLALRYLOWNGYAYTIGLKDMELDSSISEKR FAGFILQOLLEWMDISQEFLAVUNCHILLAWSGROW EKSPHEQERFFAKILLPLNQYFTINHLEAVYSGROW EKSPHEQERFFAKILLPLNQYFTINHLEAVYSGROW EKSPHEQERFAKILLPLNQYFTINHLEAVYSGROW EKSPHEQERFFAKILLPLNQYFTINHLAG KYSQARTQVKGVGQNLTYTYALLPVLTLF QHAQBIGGRODVILDDVQVSCYRILLSINSLG TIKNTYVEKLRPALGECLARLAAMPVAFLE PQLNEYNACSVYTIKSPRERAILGERYNSVERM CPDIPVLERLMADIGGLAESAGTEMPHYEE TILPMLCSYLFRWANGEGPEAPPSALPAGAPPP CTAYTSDHLINSLLONILRIVNILGIDEASWM KRILAYFAQFIVSRAFRELIGLSHFPTIGRIKKR AGKVSEEBQLAIEAKAARGGELLVRDEFS VLCRDLYALYPLLRYVDNNRAQWLTEPPS ERIFRMYGEGITYWSKSINFACGNILRIKA AGKVSEEBQLAIEAKAARGGGGLLVRDEFS VLCRDLYALYPLLRYVDNNRAQWLTEPPS ERIFRMYGEGITYWSKSINFEQNIVVQ NEINNMSFLTADNRSKMAKAGDIQSGGSDQE RTKKKRRGROFSYOYSLLVKKMILFIGDHIPTIGNILRIK KRAVPHELBYNGROFSYNGYSLTAKKMILFIGNILRIVNILF THRACMMFLESFYKAAWILTISHFEGNIPTVO NEINNMSFLTADNRSKMAKAGDIQSGGSDQE RTKKKRRGROFSYNGYSLTAKKMILFIGNILFROFWARTHINH LKHLQGKVEGSPSLWQMALYRCVFGRE DADDPEKIVRRVQEVSAVLYYLQTGHEPYKS KRAVWHELISKQRRRAVVACPRIMTINHL THRACMMFLESFYKAAWILTISHFEGNIPTVO LSKAGBQEEEBEEVEEKKPDPLHQLVLHFBRT ALTERSKLLDEDYLVMAYADIMAKSCHLEEG GENGBAEEEVVSFERKQMEKQLLLYCHFSRT ALTERSKLLDEDYLVMAYADIMAKSCHLEEG GENGBAEEEVVSFERKQMEKQLLLYCHFSRT ALTERSKLLDEDYLVMAYADIMAKSCHLEEG GENGBAEEEVVSFERKQMEKQLLLYCHFSRT ALTERSKLLDEDYLVMAYADIMAKSCHLEEG GENGBAEEEVVSFERKQMEKQLLLYCHFSRT ALTERSKLLEDGYLVMAYADDEFTQDLFRTPLQLLC EGHNNDFQNYLRTQTONTTTINICITVDYLL RQESISDFYWYYSGROWLEGGGRANWSSTIKL DAVVGFEHYYANGKAMACAGDIGGOGGBOLELKEL LDLQKDMVYNGLILLEGNVYNGMARGMY DMVSSSNYVERMILKFTOMFTATHUNGTOSTAFL DAVVGFEHYYANGKMILFTOMOTATHUNGDSQELLKEL LDLQKDMVYNGLILLEGNVYNGMARGMY	1						
MSASFYPDIKASMVLFLDRVYGIENODFILH VLIDVGRLPDMRAA ASLDTASTEMALAV NRYLCLAVI, PLITKCAPL PAGTEHRAIMVDS MLHTVYPLSRGRSI, TKAQRDVIEDCLMSLCR YRFSMLQHLRRLVFDVPLLMERAKMPLKLL TNHYERCWKYYCL, PTOWANGYTSEELHL IRKLEWGIFDSLAHKKYDPELYRMAMPCUSEELHL TRKLEWGIFDSLAHKKYDPELYRMAMPCUSEELHL TRKLEWGIFDSLAHKKYDPELYRMAMPCUSEELHL TRKLEWGIFDSLAHKKYDPELYRMAMPCUSE AIGALAPDYVDASYSSKAEKKATVDABONP DPRPVETI, NVIIPEELLDSFINKPAEYTHEK WAF DKIQNNWSYGENIDEELKTHPMLRPYKTYSE KDERIYRWPIKESL KAMIAWFAEYTHEK WAF DKIQNNWSYGENIDEELKTHPMLRPYKTYSE KDERIYRWFRESL KAMIAWFAEYTHEK WAF DKIQNNWSYGENIDELKTHPMLRPYKTYSE KDERIYRWFRESL KAMIAWFAEYTHEK WAF DKIQNNWSYGENIDELKTHPMLRPYKTYSE KDERIYRWFRESL KAMIAWFAEYTHEK WAF DKIQNNWSYGENIDELKTHPMLRPYKTYSE KDERIYRWFRESL KAMIAWFAEYTHEK WAF DKIQNNWSYGENIDELKTHPMLRPYKTYSE KDERIYRWFRESL KAMIAWFAEYTHEK WAF QELLKFLOMMYAVTTGLKDENSEKER FAFGFLQULRWMDISGEFIAHLEAVVSSGRV EKSPHEGERFFAKILLPLINQYFTNICLYFISL TRKAVLGSGGHASINGERGHANTSIFFCKLAALV RHRYSLFGTDAPAVVNCHHLARSLDARTYM KSQPEVKAGLASFYESASEDWILDSYSEKER FAFGFLQULRWMGRGGASKREKEMYENLGL KYSQAKTQVKGYGQNLTYTTVALLPLYTLLFY QHAQHQFGDDVLDDVQXSTTLCSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLE PQLNEYNACSVYTTKSPREALIGPNSVERM CPOPIPVLERLMADIGGLASSGARYTEMPHVE TILPHLCSYLFRWWERGFBASARYTEMPHVE TILPHLCSYLFRWWERGFBASAPLAGAPPP CTAYTSDHLNSLLGNIFTINGLRKR AGKVYSEEPGLALBAKARAGGGGLUNDEFS VLCRDLYALYPLLRTYDNNRAQWLTERNPS AEELFRAVGEHTYWSKSINFKREEQNFVQ NEINNMSFLTADNKSKMAKAGDIQSGOQE RTKKKRRGRYSVOYSLUATLKKMLPIGLN MCAPTTDOLITLAKTRYALKDTIDEWFREFLH NILHLGGKVGSPSLRVGMKQRLLYYLDOTSTHPYKS KKAWWHALLSKQRRAVACREMTPLYYKS KKAWWHALLSKQRRAVACREMTPLYYLL THRACAMMYLGSSCGETGAMVSSTLLL GISLINGGNAEVOYKMLDYLKDKKEVGFFQS IQALMOTCSVLOLANAFGROKKGQLLLYGERGP ALTERSKLEDEDYLVMAYADIMAKSCHLEG GENGGAEEEVEVSFEEKQMKQNLLYLLUTRSTT ALTERSKLEDEDYLVMAYADIMAKSCHLEG GENGGAEEEVEVSFEEKQMKQNLLYLLUTRSTT ALTERSKLEDEDYLVMAYADIMAKSCHLEG GENGGAEEEVEVSFEEKQMEKQNLLYQARL HTRGAAEMVLQMISACKGETGAMVSSTLLL GISLINGGNAEVOYKMLDYLKDKEVGFFQS IQALMOTCSVLVALAMAMMACALQOSQEELKEL LDLQKDMVVMLISLLEGNVVNGMARQMY DMVSSSNVEMMLKPTOMPTITUCTVDYLL LEGISDFYWYNGGMARAM	ł	Į.					
VILDUGILPMRAAASLDTATFSTIEMALAV NRYLCLAVLPITRCAPLEAGTERRAMVDS MLHTVYRLSRGRSLTKAQROVIEDCUMSLCR YRPSMLOHLIRLRIVETVOPULNERAKMPLKIL TRIKLPROMENT LIERAKMPLKIL TRIKLPROMENT LIERAKMPLKIL TRIKLPROMENT LIERAKMPLKIL TRIKLPROMENT LIERAKMPLKIL TRIKLPROMENT LIERAKMPLKIL TRIKLPROMENT LIERAKMPLKIL TRIKLPROMENT LIERAKMPLKIL TRIKLPROMENT LIERAKMPLKILSE AUGUNNEWS GENIDELIKITHPMLRPYKITSE KOKENTRWPIKCESLKAMMEWTIEKAKTOBENNE BERTEKKKTAKISOSAQTYDPREGVINOPPDL SANTLSRELQAMAEQLAENYINITWORKKQ ELERKGGTHPILLPYNTITITAKEKARDREKA QELLKFLOMNGYAYTRELDMELDSSSIEKR FARGFLQOLLE MWDISGEFALIELAVYSGRW ESPHEQEKFFAKILLPLNQYFTINHLYTIS TRAKVLGSGGARSNEKEKEMISLEVSSGRW ESPHEQEKFFAKILLPLNQYFTINHLYTIS TRAKVLGSGGARSNEKEKEMTSLPCKLAALV RHYSLFGTDAPAVVNCHHLARSLDARTVM KSOPENYAGLERSFESASEDKWENLLQ KVSQARTQVKGVGQNLTYTTVALLPVLTILF QHAQBIGRODVILDDVQVSCYRILCSINSLG TIKNTYVEKLRPALGECLARLAAMPVAFLE PQLINEYNACSYVTIKSPREARLICLEPNSVERM CPUPPVLERLMADIGGLAESGARYTEMPHVE ITLPMLCSYLFRWERGFEAPPSALFAGAPPP CTAVTSDHLINSLLONILRIVNILGIDEASWM KRILSYFAQPIVSRAFPELLGSHFPTIGRIKKR AGKVSEEBQLAIRAKARAGGGGLLVRDEFS VLCRDLYALYPILIRYVDNINRAQWLTEPNS AEELFRAVGGETYWSKSINFKREEQNFVVQ NEINNMSFLTADNISKMAKAGDIQSGGSDQE RTKKKRRGRSYSVOTSLLINKLMIPGIAN MCAPTDODLITLAKTRYALKOTDEBVREPLH NNLHLQGKVEGSPSLRWQMALYRGVGREE DADDPBKIVRRVGRSYAVLYTLOTHETPNS KRAVWHELISKQRRRAVVACPRIMTPLINLP THERACMMFLESYKAAWILLISKPEDRIMDD LSKAGBGEEBEERVEEKKPDPLHQLVLHPSRT ALTERSKLLDEDYLVLMAYADIMAKSCHLEG GENGBAEBEVSYSTEKQMEKQBLLYRLHSRIL LIGISLNGGNAEVONGMLDYLKDKKEVGFPQS IQALMOTICSVLOLINAFFGRWAGLGIGMVPB DATYPRONGRIKVMANDDEFTODLERFICLLC EGHNNDFQNYLKTQTGNTTITUILUTVDYLL RQESISDFYWYYSGKDVEEQGGRNFSKAM SVAKQVYNSLTEYQGPCTGNQQSLAHSRIL DAVKOFHINDRONYLKDKKEVGFPQS IQALMOTICSVLOLINAFFGRWAGLGIGMVPB DATYPRONGRIKVMANDSFILDLLC EGHNNDFQNYLKTORGENGRAGLGIGMVB DAVVGFHLYTARHMAMKLALODSQELLKEL LDLQKDMVVNGLISLEGGRUNDROMYNOLMAROMY DMVSSSNIVEMULKFPTOMETTICUTVDYLL RQESISDFYWYYSGKDULEFORGRONSTIKL	l l	į					l .
MLHTVYRLSRGRSLTKAQRDVIEDCLMSLCR YRPSMLQHLRRLVPDPILNERAMPILKLL TNHYERCWKYYCLPTGWANFGVTSEEELHL TRKLFWGIFDSLAHKKYDPELLYRMAMPCLC AIAGALPPD YVDASYSSKAEKKATVDABGNF DPRPVETLNVIIPEKLDSFINKFAEYTHEK WAF DKIQNNWSYGENIDEELKHMIR RYKTTSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KOZGRITONAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAW	· ·						l
MLHTVYRLSRGRSLTKAQRDVIEDCLMSLCR YRPSMLQHLRRLVPDPILNERAMPILKLL TNHYERCWKYYCLPTGWANFGVTSEEELHL TRKLFWGIFDSLAHKKYDPELLYRMAMPCLC AIAGALPPD YVDASYSSKAEKKATVDABGNF DPRPVETLNVIIPEKLDSFINKFAEYTHEK WAF DKIQNNWSYGENIDEELKHMIR RYKTTSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KDKENYRWPIKESLKAMMAWEWTIFSE KOZGRITONAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAW							
YIRPSMLQHL LRIL VPDVPIL NIEPAKMPLKIL TNIHYBGCWYYCL PTWAMPGVTSEEBLH TRKLFWGIFDSLAHKKYDPELYRMAMPCLC AIAGALPPD YVDASYSSKAEKKATVDAEGNP DPRVEILINVIIEKLDSPINKFAEYTHEK WAF DKQNNWSYGENIDEELKTHPML RPYKTFSE KDKENYRWPIKESLKAMWUTIEKAREGE EEKTEKKKTAKISQSAQTYDPREGYPROPPDL SAVTLSRELQAMAEQLAENYHNTWGRKKKQ ELEAKGGTHPLLVPYDTLTAKEKARDREKA QELLKFLQMNGYAVTRCIKDMELDSSIEKR FARGFLQQLLRWMDISQSFIAHLBAVVSGRV EKSPHEQEIKFFAKILLPLINQYFTINGLYFLS TPAKVLGSGGHASNIKEKEMITSLFCKLALV RHRVSLFGTDAPAVVNCLHILARSLDARTVM KSQPEIVKAGRESPFEASGHANNEKEMITSLFCKLALV RHRVSLFGTDAPAVVNCLHILARSLDARTVM KSQPEIVKAGRESPFEASGHEKMYENIRLG KVSQARTQVKGVGQNLTYTTVALLPVLTTLF QHAQHQFGDDVILDDVQVSCYRTLCSIYSLG TINNTYVEKLRPALGECLARLAAAMPVAFLE PQLNEYMACSVYTIKSPRERALICINSVEEM CPDIPVLERILMADIGGLAESGARYTEMPHVE IILPMLCSYLFRWWERGPEAPPSALPAGAPPP CTAYTSDHINSLLGNILRIVNNILGIBEASWM KRLAVFAQPIVSRARPELLQSHFFITGRLRKR AGKVVSEEBQLALEAKAEAQEGELVRDEFS VLCRDLYALFPLLRYVDNIRAQWLTEPNPS AEBLFRMVGEIFIYWSKSHNFKREEGNFVVQ NEINNMSFLTADNIKSKAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNIHLQGKVEGSPSLRWQMALYRGVPGREB DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACRRMTILVNLP THRACNMFLESYKAAWLITEDHSPGRMIDD LSKAGRQUEEREEVERFEKOMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL GISILNGGNAEVQGKMLDYLACKSTIKL LLGESISDFYWYYSGKDVEEQGKRNPSKAM SVAKQVYNSLTEYYGGPCTGOQSLAHSRLW DAVVGFLHYFAHMMKLAQDSSQELLKEL LDLQKDMVVMLLSILEGNVYNGMARQMV DMLYESSSVIVEMLKFTDMIFLKLKDIVGSSAF	1						·
TNHYERCWKYYCLTGWANFGYTSEERLHL TRKLFWGIPDSLAHKKYDELYRMAMPCLC AIAGALPPDYVDASYSSKAEKKATVDAEGNP DPRPVETLINVIIPEKLDSFINKFAEYTHEKWAF DKIQNINWSYGENIDEELKHPMLRPYKTTSE KDKEINTWPIKESLKAMMAWBUTEKAEGEG EEKTEKKKAKKAKIOSAGYTDPREGYNPOPPDL SAVTLSRELQAMAGQLAENYHNTWGRKKQ ELLAKGGGTIPHLVPYDTLTAKEKARDEKK QELLKFLQMNGYAVTRGLKDMELDSSIEKK FAFGFLQQLLR WMDISGEFIAHLBAVYSSGRV EKSPHEQEIKFFAKILLPLINQYFINHCLYFLS TPAKVLGSGGHASNIKEKEMITSLECKLALV RHRYSLFGTDAPAVVNCLHILARSLDARTYM KSGPEIVKAGLRSFFESASEDIEKNYHNLRJ GVAGARTYGKGGGOLTTTTVALLPVLTTLF QHIAQHOFGDDVLDDVQVSCYRTLAAAMPVAFLE PQLNEYNACSVYTTKSPRERALGLPNSVEEM CPDIPVLERLMADIGGLAESGARYTEMPHVE ITLPMLCSYLFRWWERGPEAPPSALFAGAPP CTAVTSDHLNSLLGNILRITVNNLGIBEASWM KRLAVFAGPUSVARAPHLISHFIFTIGRLRK AGKVVSEEQLALFAKAEQGGELLVRDEFS VLCRDLYALTYPLIRKYVNRAQWLTEPNPS AEELFRWGEIFTYWSKSHNFKREGNFVVQ NEINNMSFLADNIKSKAAGDIGSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAFTDQDLTILAKAADIGVGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAFTDQDLTILAKGADIGSGSODGE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAFTDQDLTILAKGADIGSGFORE STIKKGRRGDRYSVQTSLIVATLKKMLPIGLN MCAFTDQDLTILAKGADIGSGFORE GENGGAEEEVEVSFEKQMEKQRLLYQGRI LYMGRESTILASKAGREGGGNVE DADDPEKIVRRVGEVSAVLYYLLOFHPYKS KKAVWHKLISKQRRRAVVACFRMFILVNLE THRACAMFLESYKAAWULTEDHISFERMIDD LSKAGFGGEFEEEVEEKREKGRIFSKAM HTRRGAABMVLQMSACKGETGAMVSSTILKL GISILNGGNAEVQCKMLDYLAKKKRGFTGAMVSSTILKL GISILNGGNAEVQCKMLDYLAKKRAGFIGGMVNE DGTVINRQNGEKVMADDEFTODLTHILUCTUPYLL RLOESISDFYWYYSGKDVUEGQGKRNFSKAM SVAKQVFNSLTETYGGPCTGNQSLAHSRLW DAVVGFLHVFAHMMARLQDSSQELLKEL LDLQKDMVVMLLSLLEGNVYNGMARQMI		<b>.</b>					l •
TRKLFWGIFDSLAHKKYDPELYMAMPCLC AIAGALPPDYYDASYSKAEKATVDABGNP DPRPVETLNVIIPEKLDSFINKFAEYTHEKWAF DKIQNNWSYGENIDEELKTHFMLRYKTFSE KDKEIYKWPIKESLKAMIAWBWTIEKARGER EEKTEKKKTAKISQSAQTYDPREGYNPQPPDL SASTISBELQAMAEQLENYHNTWGRKKQ ELEAKGGGTHPILVPYDTITAKEKARDEKA QELLEFILQMIGYAVTTHICLYFLS TPASVLGSGGHASNEKEMITSLFCLKAALV RHRYSLFGTDAPAVVNCLHILARSLDARTYM KSGFEIVKAGGSGFFAKILLPLINQYTTHICLYFLS TPASVLGSGGHASNEKEMITSLFCLKAALV RHRYSLFGTDAPAVVNCLHILARSLDARTYM KSGFEIVKAGGRGFFAKILLPLINQYTTHICLYFLS TPASVLGGGGHASNEKEMITSLFCLKAALV RHRYSLFGTDAPAVVNCLHILARSLDARTYM KSGFEIVKAGLSFFFSASDIEKMWFAILRIG KVSQARTQVKGVGQNLTYTTVALLPVLTTLF QHAQHQFDDVLLOVYSCYRILCSIYSLG TIKNTYVEKLRPALGECLARLAAAMPVAFLE PQLNEYNACSVYTTKSFRERAILGINSVERM CPDPVLERLMADIGELASGARYTEMPHVIE ITLPMLCSYLPRWWERGFEAPPSALPGAGPF TAVTSDHINSLIGNILRIVNNLGIDEASWM KRLAYFAQPIVSRARPBLLQSHEPTIGRLRKR AGKVVSEEQLALAKARQHGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRNYGEIFTYWSKSINHFKEEQNFVVQ NEINMSSTLTADINKSKMAKAGDIOSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRALKDTDEERVEFLH NNLHLQGKVEGSPSLRWQMALYRGVFGREE DADDPEKTVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACPRMTPLYNLP THRACNMFLESYKAARWILTEDHSFEDRMIDD LSKAGGGGEEEREEVESKEFEKOPEKQHEKMID LSKAGGGGEEEREEVESKEFEKOPEKQHEKMID LSKAGGGGEEEREEVESKEFEKOPEKQHEKMID LSKAGGGGEEEREEVESKEFEKOPEKQHEKMID LSKAGGGGEEEREEVESKEFEKOPEKQHEKMID GTYNRRQNGEKVMAADEFTQDLFFFT,QLLC EGHNINDFONYLRTGTATTNIHICTVDYLL RLQSISDFYWYYSGKDVEEQGKRIFSKAM SVAKQYFNSLTEYUGGCKROFSKAMS SVAKQYFNSLTEYUGGCKRIFSKAM SVAKQYFNSLTEYUGGCKRIFSKAM SVAKQYFNSLTEYUGGCKRIFSKAM SVAKQYFNSLTEYUGGCKRIFSKAM SVAKQYFNSLTEYUGGCKRIFSKAM SVAKQYFNSLTEYUGGCKRIFSKAM SVAKQYFNSLTEYUGGCROPALKELEL LDLQKDMYWMLISLLEGNVYNOMIARQMY DMLYSSSSTVEMILKFTDMFLKKLDIVGSSAF	1	1	· .		'		l
ALAGALPPDYVDASYSSKARKKATVDABGINF DPRPVETILNVIPEKLIPNKFAEYTHEK WAF DKIQNINWSYGENIDEELKTHPMIRPYKTFSE KDKEIYRWPIKESLKAMMAWEWTIEKAREGE EEKTEKKKTAKISQSQATYDPREGYNPQPPDL SAVTLSRELQAMAEQLAENYHNTWGRKKKK ELAGGGTHPLLVPTITAKEKKARDREKA QELLKFLQMINGYAVTRGLKDMELDSSSIEKR PAFOFLQQLLRWMDISQEFIAHLEAVVSSGRV EKSPHEQEIKFFAKILLPLINQYFTHHCLYFLS TPAKVLGSGGHASNKEKEMITSJECKLAALV RHRYSLEFOTDAPAVVICHILARSLDARTVM KSGPEIVKAGLRSFFESASEDIERMVENLRLG KVSQARTQVKGVGQNLTYTTVALLPVLTTLF QHIAQHQPGDDVILDDVQVSCYRTICSIVSLG TIKNTYVEKRPALGECLARLAAMPVAFLE PQIMEYYHACSYYTIKSFERALIGLIPNSVERM CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERFALIGLISHFTIGRLKR CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERALGGEPAPPSALFAGAPPP CTAVTSDHLNSLLGNILRIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLKR AGKVVSEEQLALLAKAAQWCELLVRDEFS VLCRDLYALYPLLRYYDNINAGWLTEPNPS AEPLFRWYGEIFTYWSKSHIPKREEGNFVQ NEINIMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTTOQDLITLAKTRYALKDTIDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRVQRVSVAVLYYLDQTIEHPYKS KKAVWHKLLSKQRRRAVVACPRMTPLVNLP THRACNIMFLESYKA AWLTEDHSFEDRIMDD LSKAGEGEEEBEVSVSFERKQMERGALLYQOARL HTRGAABWJLQMISACKGETGAMYSSTILL GISLINGGMAEYOQKMLDYLKDKREVGFIQS IQALMQTCSVLDLNAFEQNKAEGLGMVNB DGTVINRQNGEKVMADDEFTQLIFFT,QLLC EGHNINDFONYLRICTOTITNINIICTVDYLL RLQSISDFYWYYGRKOMLDYLRJARGVEFRIS DGTVINRQNGEKVMADDEFTQLIFFT,QLLC EGHNINGFONYLRICTOTITNINIICTVDYLL RLQSISDFYWYYGRKOKABGLGGWKNFSKAM SVAKQVFINSLTEYUGGCKRIPSKAM SVAKQVFINSLTEYUGGCKRIPSKAM SVAKQVFINSLTEYUGGCKRIPSKAM SVAKQVFINSLTEYUGGCKRIPSKAM SVAKQVFINSLTEYUGGCKRIPSKAM SVAKQVFINSLTEYUGGCKRIPSKAM SVAKQVFINSLTEYUGGCKRIPSKAM SVAKQVFINSLTEYUGGCKRIPSKAM SVAKQVFINSLTEYUGGCKRIPSKAM SVAKQVFINSLTEYUGGCKRIPSKAM SVAKQVFINSLTEYUGGCFCORQQSALARSILW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMALISLLEGNVVNOMIARROMV DMLYSSSNVEMILKFFDMTLKKKNIVGSEAF							l
DPRPVETLNVIPEKLDSFINKFAEYTHEK WAF DKJONNWSYGENDEKTHPMLRPYKTSE KDKEIYRWPIKESLKAMIAWEWTIEKAREGE EEKTEKKET AKTISQSAQTYDPREGYNOPPDIL SAVTLSSELQAMAEQLAENYHNTWGRKKQ ELEAKGGGTHPILVPYDTITAKEKARDREKA QELLKFILQMIGYADTISLOMEDISSSIEKR FAFGFLQQLILRWMDISQEFIAHLEAVVSSGRV EKSPHEQEIKFFAKIL PINQYFTHHELYFLS TPAKVLOSGGHASNKEKEMITSLPCKLAALV RHRYSLRGTDAPAVVNCLHILARSLDARTYM KSGPEIVKKAGLRSFFESEDIEKMVENLEGI KVSQARTQVKGYGQNLTYTTVALLPVLTTLF QHIQHQFDDVILLOVYSCYRILCSIYSLG TTKNTTYVEKLRPALGECLARLAAAPPVAFLE PQLNEYNACSVYTTKSPERALIGLPNSVEBM CPDIPVLERLMADIGGLAESGARYTEMPHVBE ITLPMLCSYLPRWWERGFEAPPSALPAGAPPP CTAYTSDHLNSLLGNIHINNINGIDEASWM KRLAYFAQPIVSRARPBLLQSHFBTIGRI.RKR AGKVVSEEQLALBAKAQFGELLVRDEFS VLCRDLYALYPLLRYVDNNRAQWLTEPNPS AEELFRAVGEIFTYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIGSGGSDQE RTKKKRRGDRYSVQTSLIVALTKKMLPIGEL MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADPEKIVRVQEVSAULYJLDQTBEPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGFGGEEVENSFERKOMEGKLILVHSST ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGGAEEVENSFERKOMEKQRLLYQARL HTRGAAEMVLQMISACKGETGAMYSSTI.KL GISILNGGNAEVQQKMLDYLKDKKEVGFPQS IQALMQTCSVLDLNAFERQMKEGELGMVMB DGTVTNRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFONYLRTQTGTTTTNIHICTVDYLL RLQESISDFYWYYSGKDVEEQGKRNFSKAM SVAKOVPISLTEYURGCRORGORGORGLERGH DGYVRSTERYPROFITENINGL DAVVGFLHVFAHMMMKLAQDSSQELLKEL LDLQKDMVYMLLSLLEGNVTNOMIARROM DAVVGFLHVFAHMMMKLAQDSSQELLKEL LDLQKDMVVMLLSLLEGNVTNOMIARROM	1						l
DKIQNNWSYGENIDELKITHPMLRPYKTYSE KDKEIYRWPIKESLAMIAWBWTIEKAREGE EEKTEKKKTAKISQSAQTYDPREGYNPQPPDL SAVTLSRELQAMAEQLAENYINTWGRKKKQ ELAKGGGTHPLYDTUTTAKEKARDREKA QELLKPLQMINGYAVTRGIKDMELDSSSIEKR PAFGPLQQLLRWMISQEFIAHLEAVYSGGTV EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS TPAKVLGSGGHASNIKEEMITSLFCLAALV RHRVSLFGTDAPAVVNCLHILAREAVYSGRV KSGPEIVKAGLRSFFESASEDIEKMVENLRLG KVSQARTOVKGGPANLTYTVALLPVLTTLP QHIAQHQFGGDVILDDVQVSCYRTLCSIVSLG TIKNTYVEKLRPALGECLARLAAMPVAFLE PQLINEYNACSVYTIKSPRERAILGLPNSVEEM CPDIPVLERLMADIGGLAESGARYTEMPHYVE ITLPMLCSYLFRWERGPEAPPSALPAGAPPP CTAVTSDHLINSLLGNILRIVNNLGIDEASWM KRLAVFAQPIVSRARPFELQSHIPPTIGRLRIKR AGKVVSEEDLALAKARAQFGEGLLVRDEPS VLCRDLYALYPLLIRYVDNNRAQWLTEPINS AELFRMVGEIFIYWSKSHNFKREEDNIFVQ NEINNMSFLTADNKSKMAKAGDIGSGGSDQE RTKKKRGDRYSVGSSANLYVLDGTEHPYKS KKAVWHKILJSKQRRAAVVACFRMTPLYNLP THRACAMFLESYKAAULTEDHSPERMIDD LSKAGEGEEEVEVSFEKKMEKGRILYQQARL HTRGAAEMVLQMISACKGETGAMVSSTIKL GISLNGGNAEVQGVSSAVLYYLDGTEHPYKS KKAVWHKILJSKQRRAAVVACFRMTPLYNLP THRACAMFLESYKAAULTEDHSPERMIDD LSKAGEGEEEVEVSFEKKMEKGRLIYQQARL HTRGAAEMVLQMISACKGETGAMVSSTIKL GISLNGGNAEVQGMLDYLKDKEVGFGS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTYVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGRITTINIICTVDYLL RLQESISDFYWYYSGKDVEEGGKRNFSKAM SVAKQVFNSLTEYDGOFCTGNQGSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDKDWVNUMLSLLEGNVVNOMMARQMV DDLVJESSSNYEMERFODMFTLSLUTGGSSANEAU SDAVVNGENAMRIKLAQDSSQIELLKEL LDLQKDKDWVNOMMARQMV DDLVJESSSNYEMERFODMFTLSLUTGGSSSANEAU DDLVSSSSNYEMERFODMFTLSLUTGGSSSANEAU DDLVSSSSNYEMERFODMFTLSLUTGLAFSSALMSLU DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDKDWVNGMARQMV DDLVJESSSNYEMERFODMFTLSLUTGGSSSANEAU DDLVSSSNYEMERFODMFTLSLUTGGSGAFMFSKAM SVAKQVPNSLTEYPOMFTLKLEDGTGARMSTIKLE LDLQKDKDWVNGMARQMV DDLVJESSSNYEMERFODMFTLLSLUTGGSGAFMFSKAM SVAKQVPNSLTEYPOMFTLKLEDGTGARMSTIKLL LDLQKDKDWVNGMARQMV DDLVJESSSNYEMERFODMFTLANGRW			[		*		
EEKTEKKTAKISQSAQITVDPREGYNPOPPDL SAVTLSRELQAMAEQLAENYHNTWGRKKKQ ELAKGGGTHPILVPYDTLTAKEKARDREKA QELLKPLQMNGYAVTRGIKDMELDSSSIEKR FAFGFLQQLLRWMDISQEFIAHLEAVYSSGRV EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS TPAKVLGSGHASNIKEERMITSLFCKLAALV RHRVSLSGGHASNIKEERMITSLFCKLAALV RHRVSLSGGHASNIKEERMITSLFCKLAALV KSGPEIVKAGLRSFPESASEDIEKMVENLRLG KVSQARTQVKGVGQNLTYTTVALLPULTLF QHAQHQFGDDVLIDDVQVSCYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAMPVAFLE PQLNEYNACSVYTTKSPRERALIGLPNSVERM CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWERGPEAPPSALPAGAPPY CTAVTSDHLNSLLGNILRIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFPTTIGRLRK AGKVVSEEDCLALRAAAQAGGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMVGEIFTYWSKSHNIKREEQNFVVQ NEINNMSFILTADNKSKMAKAGDISGGSDGDDE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKUTDEEVREFLH NILHLQGKVBGSPSLWQMALVRGVPGREE DADDPEKIVRRVQEVSALYTLDQTEHPYKS KKAVWHKLLSKGRRAVYTLDQTEHPYKS KKAVWHKLLSKGRRAVYTLDQTEHPYKS KKAVMKLLSKGRRAVYTLDQTEHPYKS KKAVWHKLLSKGRRAVYTLDQTEHPYKS KKAVWHKLLSKGRRAVYTLDQTEHPYKS GGRGAEEEVEYEFKENAMAVGCHLEGG GGNGAEEEVEYEFKENAMEKGLLLVQQARL HTRGAABMVLQMISACKGETGAMVSSTIKL GISLINGGNAEVQGKMLDYLKKEVGFFQS IQALMQTCSVLDLNAFERGNKAEGLGMVNE DGTYUNRQNGEK WMADDEFTQDLLC EGHNNDFQNYLRTQTGTTTTNILICTVDYLL RLQESISDFYWYSGKDVERGQRKRESKAM SVAKQVYNSLTEYJQOPCTGNQQSLAHSRLW DAVVGFLHYPAHMMMKLAQDSSQELLKEL LDLQKDMVVMLLSLLEGNVVNGMLARQMV DMLVSSSNVEMILRFDMFLKLKDIVGGSEA			•				DKIQNNWSYGENIDEELKTHPMLRPYKTFSE
SAYTLSRELQAMAĞQLENHINTWORKKKQ ELEAKGGTHPLLVPYDILTAKEKARDREKA QELLKPLQMINGY AYTRGIKDMELDSSIEKR FAFGFI,QULRWIDISQEFIAHLEAVYSSGRY EKSPHEQEIKFFAKILLPLINQYFINHCLYFLS TPAKVLGGGGHASINEKEMİTSLFCKLAALV RHRVSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGLRSYFESASEDIEKMYENLRLG KVSQARTQVKGYGQNLTYTTVALLPVLTILF QHIAQHOFGDDLIDDVQVSYTRLCSIYSLG TTKNTYYEKLRPALGECLARLAAAMPVAFLE PQLINEYNACSVYTTKSPRERAILGLINSVERM CPDIPVLERLIMADIGGLAESGARYTEMPHIVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP CTAYTSDHLISLIGNILRIIVNINLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALBAKAAQQEGLLVRDERS VLCRDLYALYPLIRYYDNINRAQWLTEPINSS AEBLFRMYGEIFIYWSKSHNFKREEQNFVVQ NEINIMSFLTADINKSKMAKAAQDIQSGGSDQE RTKKKRRGDRYSVQTISLIVATLKKMLPİCIN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NILHLQĞKVEGSSLRWQMALYRGVPGREE DADDPEKIVRRVQFSSLRWQMALYRGVPGREE DADDPEKIVRRVQFSSLRWQMALYRGVPGREE DADDPEKIVRRVQFSSLRWOMALYRGVPGREE DADDPEKIVRRVQFSSLRWOMALYRGVPGREE DADDPEKIVRRVQFSVSAVLYYLDQTSHPYKS KKAVWHKLLSKQRRRAVVACPRMIPLYNLD TÜRRACINGHESSYKAAWILTEDHSFEDRMIDD LSKAGGQEEEEEBEVEEKKPDFLHQLVLHFSRT ALTEKSKLDEDLYMAYADIMAKSCHLEGG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAABMVLQMISACKGETGAMVSSTLKL GISILINGGNAEVQQKMLDYLKDKKEVGFFQS IQALMGTCSVLDLNAFERQNKAEGLGMYNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHINDFONYLRTQTGNTTTTINHICTVDYLL RLQGSISDFYWYYSGKDVIEQQKKRISKAM SVAKAQVFNSLTETYIQGFCTGRQQSLAHSBLW DAVVGFLHVFAHMMMKLAQDSSQELLKEL LDLQKDMVVMLLSLLEGNVYNGMARQMV	ļ						KDKEIYRWPIKESLKAMIAWEWTIEKAREGE
ELEAKGGGTHPLLVPYDTLTAKEKARDREKA QELLKFLQMNGYAVTRGLKDMELDSSSIEKR FAAGFLQQLLRWMDISQEFIAHLEAVVSSGRV EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS TPAKVLGSGGHASNKEKEMITSLPCKLAAL/ RHRVSLGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGLRSFFESASEDIEKMYENLRLG KVSQARTQVKGVGQNLTYTTVALLFVLTTLF QHAQHOFGDDVILDDVQVSCYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLE PQLNEYNACSVYTTKSPRERALGLPNSVEEM CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSVLPRWERGPEAPPSALPAGAPPP CTAVTSDHLNSLLGNILRIVNNLGIDEASWM KRLAVFAQPIVSRAPPBLLQSHTPTIGRLRKR AGKVVSEEGGLALBAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYYDNNRAQWLTEPNPS AEELFRWVGEIFTYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRVQEVSAVLYYLDQTEHPYKS KKAVPHKLLSKQRRRAVVACFMTPILYNLP TTRACNMFLESYKAAWLTEDHSFEDRMIDD LSKAGGQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEVEVSFEEKQMEKQRLLYQQARL HTRGAEMVLQMISACKGETGAMVSSTILL GISILNGGNAEVQQKMLDYLKDKKEVGFPQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKWADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEQGKRNSKAM SVAKKQVFNSLTETYQGPCTGRQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSQGLLKEL LDLQKDMVVMLLSLLEGDIVVNGMLARRUW DMVGSSSNVEMILKFTDMFLKLKDDIVGSGAF	1 .	i					EEKTEKKKTAKISQSAQTYDPREGYNPQPPDL
QELLKFLQMNGYAVTRGLKDMELDSSSIEKR FAFGFLQLLRWMDISQEFIAHILEAVVSSGRV EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS TPAKVLGSGGHASNIEKEMITSLFCKLAALV RHRVSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGLRSFFESASEDIEKMYENLRLG KVSQARTQVKGYGQNLTYTTVALLPVLTTLF QHIAQHOFGDDUDDVQVSYTRLCSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLE PQLNEYNACSVYTTKSPRERALLGLPNSVEEM CPDIPVLERLMADIGGLAESGARYTEMPHVIE TILPMLCSYLPRWWERGPEAPPSALPAGAPPP CTAVTSDHLNSLLGNILIRIVNNLGIDEASWM KRLAVFAQPIVSTSARPBILQSHFFPTIGRLKRR AGKVVSEEEQLALBAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNIRAQWLTEPNPS AEELFRMVGEIFTYWSKSHNFKREEQNFVVQ NEINNMSFLTADNIKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVGTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQBVSAVLYVLDOTHEIPYES KKAVWHKLLSKQRRRAVVACPRMTPLYNLP TTRRACNMFLESYKAA WULTEDHSFEDRMIDD LSKAGPQEEEEBEVPEKKPPPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEK(MEKQRLLVQQARL HTRGAASMYLOMISACKGETGAMVSTILKL GISILNGGNAEVQOKMLDYLKDKKEVGFFQS IQALMGTCSVLDLNAFERQNKAEGIGMYNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHINDPQNYLRTQTGRTTTINIIICTVUPVLL RLQESISDFYWYYSTSKDVEEQGKRNFSKAM SVAKQVFNSLTEYIQGFCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFTDMFLKLKDDVGSGAF		1 1				'	SAVTLSRELQAMAEQLAENYHNTWGRKKKQ
FAFGFLQQLLRWMDISQEFIAHLBAVVSGRV EKSPHEQEIKFFAKILLPLINQYFINHCLYFLS TPAKVLGSGGHASNKEKEMITSLFCKLAALV RHRYSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGLRSFPESASEDIEKMYENIRLG KVSQARTQVKGVGQNLTYTTVALLPVLTTLF QHAQHQFGDDVLLDDVQVSCYRTLCSIYSLG TTKNTYVEEKLRPALGECLARLAAMPVAFLE PQLMEYNACSVYTTKSPRERALLGLPNSVEEM CPDIPVLERLMADIGGIAESGARYTEMPHVIE ITILPMLCSYLTPRWWERGPEAPPSALPAGAPPP CTAYTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPEILQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDERS VLCRDLYALYPLLIREYVDNNRAQWLTEPNPS AEELFRNVGEITYWSKSINFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTDSLIVATLKKMLPIGLN MCAPTDQDLJTLAKTRYALKUTDIEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLP THRACNMFLESYKAAWILTEDHISFEDRMIDD LSKAGEQEEPEEVPEKKPDPLHQLVHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEREKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFRRQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIUCTVDYLL RLQESISDFYWYYSGKOVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELKEL LDLQKDMVVMLLSLLEGNVVNGMARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF		<b>j</b>					ELEAKGGGTHPLLVPYDTLTAKEKARDREKA
EKSPHEGEIKFFAKILLPLINQYFTINHCLYFLS TPAKVLGSGGHASNKEKEMITSLFCKLAALV RHRYSLGTIDAFAVVNCLHILARSLDARTVM KSGPEIVKAGLRSPFESASEDIEKMYENLRLG KVSQARTQVKGVGQNLTYTTVALLPVLTTLF QHAQHQFGDDVLLDDVQVSCYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLE PQLNEYNACSVYTTKSPRERALIGLPRSVEEM CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGFEAPPSALPAGAPPP CTAYTSDHLNSLLGINLRIVINLIGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALTYPLLIRYVDNINAQWLTEPIPS AEELFRMVGEIFIYWSKSHNFKREEQNFVQQ NEINNMSFLTADINKSKMAKAGDIQSGGSDQE RTKKKRRGBPYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTIDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTISHPYKS KKAVWHKLLSKQRRRAVVACPRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGBQREEEBEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNB DGTVTNRQNGEKVMADDEFTQDLFRFLQLLC EGHNDFQNYLRTQTGTTTTNNIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMLARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF							QELLKFLQMNGYAVTRGLKDMELDSSSIEKR
TPAKVLÖSGGHASNKEKEMİTSILPCKLAALV RHRVSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGLRSFFESASEDIEKMVENIRLG KVSQARTQVKGVGQNLTYTIVALLPVLTTLF QHIAQHQFGDDVILDVQVSCYRTICSIYSLG TITKNTYVEKLRPALGECLARLAAAMPVAFLE PQLNEYNACSVYTTKSPRERAILGLPNSVEEM CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPF CTAVTSDHLNSLLGNILRIVNNLGIDEASWM KRLAVPAQPIVSRARPELLQSHFIPTIGRILKK AGKVVSEEQLALBAKAAQBGGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRWGEIFTYWSKSHNFKREEQNFVVQ NEINNMSFLTADINKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRAVVACPRMTPLYNLP THRACNMFLSSYKAAWILTEDHSFEDRMIDD LSKAGBQEEEBEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLTMAYADIMAKSCHLEEG GENGFAEEEVEVSFEEKOMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMOTCSVLDLNAFRGONKAEGLIGMYNE DGTVINRQNGEKVMADDEFTQDLFRFLQLC EGHNNDFQNYLRTQTGTTTTNIIICTYDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGFCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVSSSSNVEMILKFPDMFLKLKDIVGSEAF		· ·					FAFGFLQQLLRWMDISQEFIAHLEAVVSSGRV
RHRVSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGLRSFFESASEDIEKMVENIRLG KVSQARTQVKGYGQNLTYTTVALLPVLTTLF QHIAQHQFGDDVIL.DDVQVSCYRTI.CSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLE PQLNEYNACSVYTTKSPRERAILGLENSVEEM CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRI.RKR AGKVVSEEQLALEAKAEAQEGLLVRDEFS VICADLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRWGEIFJYWSKSHNFKREQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRAVVACPRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGBQEEREBVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVSFFEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGMTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKOVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMUVSSSSNVEMILKFPDMFLKLKDIVGSEAF		]	]				EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS
KSGPEIVKAGLRSFFESASEDIEKMVENLRLG KVSQARTQVKGYQNLTYTITVALLPVLTTLF QHIAQHGGDDVILDDVQVSCYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLE PQLNEYNACSVYTIKSPRERAILGLPNSVEEM CPDIPVLERLMADIGGLAESGARYTEMPHIVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHIPPITGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLRYVDNNRAQWLTEPPNS AEELFRMYGEIFTYWSKSSHIPKREEQNEVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGBQEEEEBEVEKKPDPLHQLVLHFSRT ALTEKSKLDEDYYLYMAYADIMAKSCHLEEG GENGEAEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKEVGFPQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGMTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEGQGKRNPSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKSMDVYMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFPDMFLKLKDIVGSEAF	İ	1					TPAKVLGSGGHASNKEKEMITSLFCKLAALV
KVSQARTQVKGVGQNLTYTTVALLPVLTTLF QHIAQHQFGDVLDDVQSCYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLE PQLNEYNACSVYTTKSPRERAILGLPNSVEEM CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP CTAVTSDHLNSLLGNILRIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFPTTGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VICRDLYALYPLLIRYVDNNRAQWLTEPNSS AEELFRMVGEIFIYWSKSHNFKREQNFVVQ NEINNMSFLTADNKSKMAKAGDIOSGGSDQE RTKKKRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDIDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRAAVVACPRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEBEEVEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMY DMLVESSSNVEMILKFFDMFLKILKDIVSGEAF	<b>]</b> .	}					RHRVSLFGTDAPAVVNCLHILARSLDARTVM
QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAMPVAFLE PQLNEVNACSVYTTKSPRERAILGLPNSVERM CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPPSALFAGAPPP CTAVTSDHLNSLLGNILRIUVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALFAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRWGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRAVVACFRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGBQEEEEEVEEKEPPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEVEVSFEEKQMEKQRLLYQQARL HTRGAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS [QALMQTCSVLDLNAFERQNKAEGLGMVNE] DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDPQNYLTRTQTGNTTTTNIICTTOYJLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKIKDIVSGEAP	1		ŀ				KSGPEIVKAGLRSFFESASEDIEKMVENLRLG
TTKNTYVEKLRPALGECLARLAAAMPVAFLE PQLNEYNACSVYTTKSPRERAILGLPNSVEEM CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP CTAYTSDHLNSILGNILRIVNNILGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEEQLALBAKAEAQEGELLVRDEFS VICRDLYALYPLLIRYVDNNRAQWLTEPNPS AEBLFRMVGEIFTYWSKSHNFKREEQNFVVQ NEINNMSTLTADINKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACPRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNB DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RIQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMY DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF		<b>,</b>	-				
PQLNEYNACSVYTTKSPRERAILGLPNSVEEM CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP CTAVTSDHLINSLLGNILRIVNNILGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEBQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRWGEIFTYWSKSHNFKREBQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATTLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEEVEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVESFEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFFLQLLC EGHNNDFQNYLRTQTGNTTINHIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQEELLKEL LDLQKDMVVMLLSLLEGNVVNGMLARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	ŀ						, , ,
CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYYDNNRAQWLTEPNPS AEELFRMVGEIFTYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTIDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEZEEVEVEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTYVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMILSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFTDMFLKLKDIVGSEAF							
ITILPMLCSYLPRWWERGPEAPPSALPAGAPPP CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMYGEIFIYWSKSHNFKREEQNFVVQ NEINMSFLTADNKSKMAKAGDIQSGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACPRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMKLAQDSSQIELIKEL LDLQKDMVVMLLSLLEGNVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF		:					l7
CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPPPS AEELFRMVGEIFTYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACPRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHINDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELKEL LDLQKDMVVMLLSLLEGNVNGMARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	1						
KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF							
AGKVVSEEQLALEAKAEAQEGELLVRDEFS VI.CRDLYALYPLIRYVDNNRAQWLTEPNPS AEEIFMVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACPRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTILKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMY DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	1	j .					
VLCRDLYALÝPLLIRYVDNNRAQWLTEPNPS AEBLFRMVGEIFTYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREB DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	1				-		
AEELFRMVGEIFTYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF							
NEINNMSFLTADNKSKMAKAGDIQSGGSDQE RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDIDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLP THRACNMFLESYKAA WILTEDHSFEDRMIDD LSKAGEQEEEEBEVEEKRPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	1	[					l
RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACPRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	1						
MCAPTDQDLITLAKTRYALKDIDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQIEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	1						
NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACPRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDPQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF		<b>.</b> .					
DADDPEKIVRRVQEVSAVLYYLDQIEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLIKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSQIELLKEL LDLQKDMVVMLLSILLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	[			i i			
KKAVWHKLLSKQRRRAVVACFRMTPLYNLP THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFFFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	1	j.					
THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	i						
LSKAGEQEEEEEVEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHINNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF							
ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	1 .	]	[				
GENGEAEEVEVSFEEKQMEKQRLLYQQARL HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	[ .	· :			İ		· · · · · · · · · · · · · · · · · · ·
HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	1						
GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESIDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	1						
IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	1						
DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	]	J '	J I			,	
EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF							
RLQESISDFYWYYSĞKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF		1					
SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF							
DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	ľ						
LDLQKDMVVMLLSLLEGNVVNGMIARQMV DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF	1	]					
DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF							
	1						
Land Anna Anna Anna Anna Anna Anna Anna A							
		L				L	An I A Introprievent Avenue Avenue I Av

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide	f	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		]	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		i		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ŀ	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1	1	peptide		/=possible nucleotide deletion, \=possible
<u> </u>		<u> </u>		sequence		nucleotide insertion
						QFLLSCSEADENEMINCEEFANRFQEPARDIG
						FNVAVLLTNLSEHVPHDPRLHNFLELAESILE
1		·				YFRPYLGRIEIMGASRRIERIYFEISETNRAQW
1 .						EMPQVKESKRQFIFDVVNEGGEAEKMELFVS
						FCEDTIFEMQIAAQISEPEGEPETDEDEGAGA
1						AEAGAEGAEGAAGLEGTAATAAAGATARV
						VAAAGRALRGLSYRSLRRRVRRLRRLTAREA
]						ATAVAALLWAAVTRAGAAGAGAAAGALGL
						LWGSLFGGGLVEGAKKVTVTELLAGMPDPT
1			1			SDEVHGEQPAGPGGDADGEGASEGAGDAAE
1						GAGDEEEAVHEAGPGGADGAVAVTDGGPFR PEGAGGI GDMGDYTPAEDDTBEGSBII KRVI G
						PEGAGGLGDMGDTTPAEPPTPEGSPILKRKLG VDGVEEELPPEPEPEPEPELEPEKADAENGEK
						EEVPEPTPEPPKKQAPPSPPPKKEEAGGEFWG
						ELEVQRVKFLNYLSRNFYTLRFLALFLAFAIN
				·		FILLFYKVSDSPPGEDDMEGSAAGDVSGAGS
		ĺ		İ		GGSSGWGLGAGEEAEGDEDENMVYYFLEES
						TGYMEPALRCLSLLHTLVAFLCIIGYNCLKVP
						LVIFKREKELARKLEFDGLYITEQPEDDDVKG
J J	ļ		]			QWDRLVLNTPSFPSNYWDKFVKRKVLDKHG
l i		ĺ				DIYGRERIAELLGMDLATLEITAHNERKPNPP
1						PGLLTWLMSIDVKYQIWKFGVIFTDNSFLYLG
1						WYMVMSLLGHYNNFFFAAHLLDIAMGVKTL
1						RTILSSVTHNGKQLVMTVGLLAVVVYLYTVV
1		I	1			AFNFFRKFYNKSEDEDEPDMKCDDMMTCYL
1 1			Ì	•		FHMYVGVRAGGGIGDEIEDPAGDEYELYRVV
1						FDITFFFFVIVILLAIIQGLIIDAFGELRDQQEQV
ļ †				1		KEDMETKCFICGIGSDYFDTTPHGFETHTLEE
		- 1	1	i		HNLANYMFFLMYLINKDETEHTGQESYVWK
501	1851	$\overline{\mathbf{A}}$	3869	467	665	MYQERCWDFFPAGDCFRKQYEDQLS
301	ומו	^	3009	467	000	VIVAIYCQLIFDKGAKTIQ*PFQQIAL/CKRMK
	•	j	- 1			LGPCFTPCGKINSEWIRELSVRVKTIKHLEIGV N
502	1852	A	3888	1042	724	
502	1002	^	3000	1042	124	SGMQWRDLTPLQPLPPRFKQFSCLSLPGSWD
1 1	i	- 1	- 1	ł	1	YRHAP\PLLTNF\*FLVEMGFCYVGQAGRKLL
1 1						ASSDQSALASQSAGITGISTAPGPPFFFLNFEA GSCSVAQAGVQ
503	1853	A	3891	1773	1193	
]		I			- 1/3	EVDSQSGVQ*QAPGSLQLQTPGLK/VSCLLSR QDYRSSLPHLASCCYYYYYY/VFL*RRGLTTL
]	ļ			ł		VQGGLKLLPSSNPFASAP*TAGITGMSHCAGP
	ł	- 1	.	į		HFNF*MFRKISCIRE*F*HTRIYDIPFLILFFKET
	į	ł		j	1	WVLLCYPGWPQIPGLKPSSCLRLLSSWDHRC
		l	].	İ	l	APPCPASFFIFHVDRVSPPCPGLVSITFKMLLL
			1	1	ľ	L
504	1854	В	3896	279	70	MVSKSKSILMSYNHVELTFSDMKKMPEAFRR
		0		· -		TQKHTTYLIPYQVIFWSTGKDAMRSFMMPFY
1		ł	1	l	i	OKEYYENO*
505	1855	A	3899	2	1396	EPGVPTKKTWFDKPDFNRTNSPGFQKKVQFG
1		-		•		NENTKLELRKVPPELNNISKLNEHFSRFGTLV
				ŀ		NLQVAYNGDPEGALIQFATYEEAKKAISSTEA
		1	ł			VLNNRFIKVYWHREGSTQQLQTTSPKVMQPL
	1	0.0	- 1		ł	VOOPILPVVKOSVKERLGPVPSSTIEPAEAOS
			1		l	ASSDLPQVLST\LLA*QKQCIIQLL/WKAAQKT
	1		1			LLVSTSAVDNNEAQKKKOEALKLOODVRKR
				ŀ		KQEILEKHIETQKMLISKLEKNKTMKSEDKAE
	1					IMKTLEVLTKNITKLKDEVKAASPGRCLPKSI
	1	İ	ŀ	- 1	• }	KTKTQMQKELLDTELDLYKKMQAGEEVTEL
	•	l	, [	İ	ł	RRKYTELQLEAAKRGILSSGRGRGIHSRGRGA
<u> </u>		- 1		ļ		VHGRGRGRGRGVPGHAVVDHRPRALEIS
					L	

WO 01/57188

PCT/US01/03800

SEQ ID	SEO ID	1174	Leno	1 10 10 1	180.00	T
NO: of	NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in in	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
cotide	seq-		USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence	1	09/496	correspondi	to last amino	l=Isoleucine, K=Lysine, L=Leucine,
neuce		}	914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline,
401.00	ľ		717	amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l .		peptide	Sequence	
		f		sequence		nucleotide insertion
	<del> </del> -		┼	sequence	<del> </del>	
			]			AFTESDREDLLPHFAQYGEIEDCQIDDSSLHA
		1				VITEKTRAEAEAAAVHGARFKGQDLKLAWN
506	1856	A	3911	1952	919	KPVTNISAVETEEVEPDEEEQREIIIA
500	1050	^	3311	1932	919	DAELSGTLSLVLTQCCKRIKDTVQKLASDHK DIHSSVSRVGKAIDKNFDSDISSVGIDGCWQA
	]	l				
		1	]			DSQRLLNEVMVEHFFRQGMLDVAEELCQES
			1	1		GLSVDPSQKEPFVELNRILEALKVRVLRPALE
	]	ľ	İ	l	ł	WAVSNREMLIAQNSSLEFKLHRLYFISLLMG
	l		1			GTTNQREALQYAKNFQPFALNHQKDIQVLM
						GSLYYLRQGIENSPYVHLLDANQWADICDIFT
			!	1		RDACALLGLSVESPLSVSFSAGCVALPALINIK
	1	}		j		AVIEQROCTGVWNQKDELPIEV\DLG*KSAGY
	· .					HSIFACPILRQQTTDNNPPMKLVCGHIISRDAL
507	1857	A	3936	439	18	NKMFNGSKLKCPYCPMEQSPGDAKQIFF
30,	1007	1	3330	439	10	SHPFSPAPGICPDAPPPLPRPSKGLGHPGTAGA
	ĺ	ĺ	1			PGSGARCHPPSTCSPSWASPG*GAKASPALPR
				·		SHGVTLLCKAQAHLCRGEDSKDASGSTSQA
	l	Ì		·		WEPG*GAWGMPRCQGPALGSCFCPPGTTVQ
508	1858	A	3944	120	412	RPAKQRDKRNRHLGR
300	1038	^	3944	120	412	WCPAGTLDFPGPQEMVLLEIEVMNQLNHRNL
	١.					IQLYAAIETPHEIVLFME\YECPK*W*GLGGGT
						TRHGASRGGVCAHSIEGGELFERIVDEDYHLT
509	1859	A	3949	31	392	EV
307	1655	Ι^	3747	31	392	LTKTPSPREKGRGVLSVLLMMI*KCRVIFVKIP
	Ì	ľ	1			MVFFLQNFC/RILLNVA\WTGD*PNTL*KEQRG
			j .			ITFSDSKS*YKATKIKTMWYCHKNRYID/ERN
510	1860	A	3954	1013	885	RIEIPEINPCICDKIIFRKLSMTTQ
	1000	Α	3934	1013	003	FSETRACCPRLEHSGRIEAHCSLNIPGSSDPPT
511	1861	Α	3956	1	1054	SASSVAATTG
"	1001	Λ.	3730	•	1034	PPAWAPRSPLIWAPTSGRHPCRAALPWSTSSV
			]			RWQPSEKQPPPPAHRGPADSLSTAAGAAELS
]	]		]		}	AEGAGKSRGSGEQDWVNRPKTVRDTLLALH
			i			QHGHSGPFESKFKKEPALTAVARTARKRKPS
						PEPEGEVGPPK\TTERPSRGCPHPQRGSRSP*L LHPLLCLRHHPLPHLIPTGPHRLKRPRM\P\SP
						MAALILVADNAGGSHASKDANQVHSTTRRN
						SNSPPSPSSMNQRRLGPREVGGQGAGNTGGL
ł			1			EPVHPASLPDSSLATSAPLCCTLCHERLEDTH
						FVQCPSVPSHKFCFPCSRQSIKQQGASGEVYC
			!			PSGEKCPLVGSNVPWAFMQGEIATILAGDVK
						VKKERDS
512	1862	Α	3957	1086	3	QDRARLDCSSATSAHCNLRLPGS*DSPASASR
		•	""	2000	-	VAGITOTHHHTWLILGSSVQTGFDHVGQAG
j				ı	l	LELLTSGDPPISASESAGIMGMSHCVWP*SWG
ļ				İ		LSHHMAPPQGDGGRARGTPGPEQSFWNLSC
			1 1		1	H*PRCQVPS*LMTQL/FWGRHQYNPTMKRGK
						LRHREACSLPLPGEGEPGLQPSS\*SQNPCSSPL
					*	FHHGL*AWLWCPELLLQGQARRH*RSPPS/FK
1				ļ .	Į.	
. 1					ĺ	CPATLSLTAWSQTKRLRSQFLLLPWL*RAL*H PP\CHWPSRRSLGDPLLPRSQG*RDGT*ASTFC
ł				Í	1	SYF*DTESHLVAQAGVOWRDLGSLOPPCPRL
i					•	
				-	ļ	K\RFSRLSPPSSYTHRYVPSHLAESCISSRDRIP
513	1863	Α	3961	3038	A726	PSRPDRSRNSNSLSR
713	1000	<b>^</b>	2201	מכטב	476	VALTTSMCCNKQVIVIDKIKSASIADRCGALH
				L L		VGDHILSIDGTSMEYCTLAEATQFLANTTDQ
٠				ľ		THE DISTORDS AT INCORPORATE OF THE PARTY OF
				ľ		VKLEILPHHQTRLALKGPDHVKIQRSDROLT
						VKLEILPHHQTRLALKGPDHVKIQRSDRQLT WDSWASNHSSLHTNHHYNTYHPDHCRVPAL
						VKLEILPHHQTRLALKGPDHVKIQRSDROLT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO;	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ļ	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l	ì			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	i ·	1		sequence		/=possible nucleotide deletion, \=possible
		<del> </del> -		sequence		nucleotide insertion
1				i		SLASSTVGLAGQVVHTETTEVVLTADPVTGF
	-			i		GIQLQGSVFATETLSSPPLISYIEADSPAERCG VLQIGDRVMAINGIPTEDSTFEEASQLLRDSSI
				*		TSKVTLEIEFDVAESVIPSSGTFHVKLPKKHN
	ļ					VELGITISSPSSRKPGDPLVISDIKKGSVAHRT
<u> </u>	ł	(				GTLELGDKLLAIDNIRLDNCSMEDAVQILQQC
1	ì					EDLVKLKIRKDEDNSDEQESSGAIIYTVELKR
					İ	YGGPLG\TTSGTEEP\FDL*IISSLTKGGLAERT
						GAIHIGDRIL\AINSSSLKGKPLSEAIHLLQMAG
						ETVTLKIKKQTDAQSASSPKKFPISSHLSDLGD
						VEEDSSPAQKPGKLSDMYPSHGCPSVDSAVD
						SWDGSA\UDTS\YGTEGT\SFQASGY\NFNTYD
		]				WRSPKQRGS\LSPVT\KPRSQTYPDVGLSYED
				·		WDRSTASGFAGAA\DSAETEQEENFWSQALE
}						DLETCGQSGILRELEATIMSGSTMSLNHEAPT  DDGDAGGDDGGGGDGGGGGGGGGGGGGGGGGGGGGGG
						PRSPAGSDRPSFQERSSSRPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSPTPVELHKVT
						LYKDSDMEDFGFSVADGLLEKGVYVKNIRPA
						GPGDLGGLKPYDRLLQVNHVRTRDFDCCLV
				·		VPLIAESGNKLDLVISRNPLASQKSIDQQSLPG
						D*SEQNSAFFQQPSHGGNLETREPTNTL
514	1864	Α	3967	833	800	LEKQGVSGMATKRLARQLGLIRRKSIAPANG
						NLGRSKSKQLFDYLIVIDFESTCWNDGKHHH
1 1		i	·			SQEITEFPAVLLNTSTGQIDSEFQAYVQPQEHPI
			•			LSEFCMELTGIKQAQVDEGVPLKICLSQFCK
				•		WIHKIQQQKNIIFATGISEPS/DF*SKIMCICYL
515	1865		3969	400	100	VR*RISYTY*SKHKSKGC
313	1000	Α	3909	492	182	CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL
						DLLGPEFLARKPLFKTKTYQSTF*SISKNE/FTC
1				:		PNFIIEEGTDLIF\*QVKHNPCHRLTPEEGTVQL NRADS
516	1866	A	3977	2	1357	KMLC/OKESNYIRLKRAKMDKSMFVKIKTLGI
				-	1337	GAFGEVCLARKVDTKALYATKTLRKKDVLL
١. ا			}			RNOVAHVKAERDILAEADNEWVVRLYYSFO
) }			]		J	DKDNLYFVMDYIPGGDMMSLLIRMGIFPESL
					*	ARFYLAELTCAVESVHKMGFIHRDIKPDNILID
1 1			ļ	ı	ļ	RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP
				1		RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA
					}	RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL
			I	l	ł	CDWWSVGVILFEMLVGQPPFLAQTPLETQM
						KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE
						DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS
			ì			AFKQFP*NHTTPTDTSNFDP\VDPDKLWSDDN
		l		1		EEENVNDTLNGWYKNGKHPEHAFYEFTFRRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD
				· [		QNTGSEIKNRDLVYV
517	1867	$\overline{\mathbf{A}}$	3980	1358	1022	FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY
J 1					- 1	VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD
	1		ĺ	ĺ	İ	*RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA
Į J		J	I			OAIFOPOPPKVLGLOV
518	1868	A	3986	974	666	SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F
						SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH
	j	J				VGQAGLELLTSGDLPALASQSAGITG\SHRAR
				. (		PENGFENIF
519	1869	A	3994	751	126	NOGLRHVGLCRTCLVNOMFASSILGKSHIHIS
	l				. ]	LISINQGHNALWKAAG\PLPLKAGYC\OSFSPC
[		ſ	ı	i	i	DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS
1	ļ	j	J	1	J	QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV
L [		İ				HLFSSEMGE\NRPMVG\ARHVYSNAALLSFTP

SEQ ID	SEQ ID	Met	Lero	N-32-4-4	Designation of	
NO: of	NO: of	hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine O=Cysteine,
nucl-	peptide	l mod	in NO:	beginning nucleotide	nucleotide	D-Aspartic Acid, E-Ghrtamic Acid,
cotide	seq-	1	USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, 1=Isoleucine, K=Lysine, L=Leucine,
seq-	иепсе	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ŀ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	}	J	114	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l		peptide	Sequence	possible nucleotide deletion, possible
	1	1		sequence		nucleotide insertion
			<del> </del>	204201100		LRCLGGEKHKSGLHARPVIVPSLELHYDMDS1
	1		}		1	AHV\FADLLLIITLPSYYIPFC
520	1870	A	3999	882	698	QSFRLSLLSSWDYRHM*PRLANF*TVFFCRDR/
_		**			] ""	SLALLPRLVSNSWPQAILPPRPPKVLGLQT
521	1871	Α	4011	1346	1178	FFF*ETVSCSAS*AGVRSHDNSSLQPPSPG\SSN
		ı				PPTSASHVAGATGTHHHAWLLSV
522	1872	A	4015	2	377	QGIALLTRMGESVKHVTGGYKLRTRPLEFAA
	1		1		1	IGDYLDTFALKLGTIDRIAQRIIKEEIEYLVELR
		Į.			1	EYGPVYSTWSALEGELAEPLEGVSACIGNCST
	l					AL*ELTDDMTEDFLFVLREYILYSDSMK
523	1873	Α	4018	341	19	ERVIHNQIQQAQRSPHIFNARRSS/PRPNIVELP
	1	İ				KVKEVCKTSKS/GQVIYKGVSIRLRANFLAEP
		i			1	L*NRREWDEAIKVLKEKQ\FLSKMVYPANLSF
		l				GNEGDITSFPAK
524	1874	Α .	4020	1067	743	FFLRWSL/DSVAQAGVKWCNLGSLQAPPPGF
		ł				TPFSCLSLPSSWDYRHPPPRLAN*LTNFLCF**
		1				RQGFTVLARMVLIS*PHDLPASASQSAGITGL
		<u> </u>				SHCSWPTSSILS
525	1875	Α	4021	781	351	QFRVIFFFLRRSHSVAQAGMQWHDHSLLQPL
						PPRLKQ/F/SHLSPPSIWDYRRVPPCLVNFSIFF
		l				VETGSCQPCLQLLGSSNPPASASQSAGIAGISH
						QGQPE*SFDIRFACVIAALRETFQCLCSASRVN
504	1000					NKIINRPTHPVESSF
526	1876	Α	4024	80	341	TPSSTSRGTEEQQSSKMAWQRREEKEHLNVR
						RSSAEDGWKADKP/VDG*TPGEDHLPTPSPFQ
527	1877	Α	4006	502	000	LHIHSSESQLHHSVKSPPSLSFRLM
321	10//	A	4026	593	230	DFYLYPERKKRGQMMTAVSLTTRPQESVAFE
				·		DVAVYFTTKEWAIMG\PAERALYRDYMLEN
				j		YGGCGPL*CHPTSKPALVFS\LEQGKESCFSPA
528	1878	Α	4028	1160	242	TGSSLSRNDWRAGWIGYLELRRYTYLS GTSELLCIQRWNWGPAFPPRPGLALAPTLOLL
	1070	Λ	1026	1100	242	VEMGSAKSVPVTPARPPPHNKHLARVADPRS
						PSAGILRTPIQVESSPQPGLPAGEQLEGLKHAO
				1		DSDPRSPTLGIARTPMKTSSGDPPSPLVKOLSE
						VFETEDSKSNLPPEPVLPPEAPLSSELDLPLGT
						QLSVEEQMPPWNQTEFPSKQVFSKEEARQPT
				1		ETPVASQSSDKPSRDPETPRSS\GSMRNRWKP\
				ļ		NSSKVL\GKSPLHPSCQDDNSPGTLTLRQGKA
						AFKPLSENVSELK/EGA/ILGTGR/LLKTEGRA
i						WEQGQD\HDKENQHFPLVES
529	1879	A	4039	2	366	KDMVLIMEMQSMITMKCPQYL*E*RKIPDITK
1			-,			CW*GCGSTGILIFC/WS*PL*KTI*OPR*FKOI*T
1					. ]	ILTIIYSIM*EHTFHNAGV*LSDIYPRFMKGYV
				1		HTEICT*MFIAVLFVVVKTWKQF
530	1880	Α	4057	358	3	LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL
J				1	1	RKYGSRINLLKSKHDKNICTENYKT*MKEIEA
1		' I		I	į	/DTDKWKDILCSWIRRIHMKDILCSWIGRTHY
				l		VKISILPKVNYRFYLISIKIIMAI
531	1881	Α	4061	50	278	TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY
		l				HKATVIKMVWYWHRO*KFSKN/RIESSEIEPH
						IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF
		l	ı	J	]	T*KR
532	1882	Α	4069	19	368	NDLLENFKFWE*FKE*LENINGTVTEKETGGV
				ŀ		YKELSSPKYSGTRQFYGQTISNFPGKIISMVY
			1	į	į	KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL
	l			· 1	•	QIWMPVSLMNIVTLKCPT
533	1883	Α	4076	1	355	PIRKFTKVAG*KSNTPK*LAFLHINNEOFENKI/
533	1883	Α	4076	1	355	PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/ ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR
533	1883	Α	4076	1	355	

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	non	in NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
'		1		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion IFNAIPIKMPMMCMAKIEKNSS
534	1884	A	4088	3	1931	IIDSSTRRMESERSPLYRQLIDLGYLSSHWNC
) 334	1004	Α	4000	3	1931	GAPGODTKAQSMLVEQSEKLRHLSTFSHQVL
		<b>l</b> '		_		OTRLVDAAKALNLVHCHCLDIFINOAFDMOR
						DLQTTPKRLEYTRKKENELYESLMNIANRKQE
1						EMKDMIVETLNTMKEELLDDATNMEFKDVI
i l						VPENGEPVGTREIKCCIRQIQELIISRLNQAVA
ļ	İ					NKLISSVDYLRESFVGTLERCLQSLEKSQDVS
1		· '				VHITSNYLKQILNAAYHVEVTFHSGSSVTRM
		Ì				LWEQIKQIIQRITWVSPPAITLEWKRKVAQEAI
						ESLSASKLAKSICSQFRTRLNSSHEAFAASLRQ
) !		ļ				LEAGHSGRLEKTEDLWLRVRKDHAPRLARLS LESRSLODVLLHRKPKLGOELGRGOYGVVYL
						CDNWGGHFPCALKSVVPPDEKHWNDLALEF
						HYMRSLPKHERLVDLHGSVIDYNYGGGSSIA
		ļ				VLLIMERLHRDLYTGLKAGLTLETRLQIALDV
		İ				VEGIRFLHSQGLVHRDIKLKNVLLDKQNRAKI
<b>!</b>						TDLGFCKPEAMMSGSIVGTPIHMAPELFTGK
						YDNSVDVYAFGILFWYICSGSVKLPEAFERCA
		}				SKDHLWNNVRRGARPERLPVFDEECWQLME ACWDGDPLKRPLLGIVOPMLOGIMNRLCKS\
						NSEQPNRGLDDST
535	1885	A	4090	2	417	ALMPHEANYEEIFLKTDKDMDGFESGLEVRE
	1000		1000	_		IFLKTR/GLPSTLLAHIWALCDSKDCGKLSKD
]		]				HFALAFHLIT\QKLIKGIDPPLVLTPEKISPSNR
1					•	ASLQKVTELTRKPVCIIFKGTILWRITDSIWMK
	1554					HNRKRIWLRA
536	1886	Α	4102	569	829	DHQK*KNIPCSWIGRINIVKMSILPKAIYRFSAI
						PIKIPMTFFTEI*S*NVYRTTKTQE*AKAILSKK   EONLEESHYLDFK*YYRAV
537	1887	A	4104	54	281	SIDCEHLIRRMLVLDPSKRLTIAOIKEHKWML
***	100.	ļ <b>'</b> '		• •	201	IEVPVQRPVLYPQEQENEPSIGEFNEQVLRLM
		}				HSLGIDQQKTIE
538	1888	Α	4109	141	314	IRHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL
						VPRENITAWMNAIGLIITALPVS
539	1889	Α	4111	268	1	ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM
		1				I*KFAT\KKSPGPYRFTAEFSHTFKEDLVPILW
540	1000		4140	100	00/4	PLFPKIYREGILPHSFYEASITL PEPCA CRAATTIVON ENTROPEGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
340	1890	A	4142	198	2064	PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR
ļ .			-00			VRGHGGVRESGRAPQQPGRRRGRRPRKRPR
						GRWRREGCGAGGRGVCVAAWSQRSIAGNN
1	1	[				DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH
,						1GALLIGEEYGDVTFVVEKKRFPAHRVILAAR
] .	1	1				CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT
[ ·		1	·			MLLKYTYTGRATLTDEKEEVLLDFLSLAHKY
		1			İ	GFPELEDSTSEYLCTILNIQNVCMTFDVASLY
	1					SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK
		ļ	•			TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK   ENHAEIMOAVRLPLMSLTELLNVVRPSGLLSP
[	[	1				DAILDAIKVRSESRDMDLNYRGMLIPEENIAT
}	ļ					MKYGAQVVKGELKSALLDGDTQNYDLDHG
						FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR
		}				DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS
1		1	-		.	WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF
<b>!</b>			!			ECMFTNKTFILEKGLIVPMENVATIADCASVI
	}		İ			EGVSRSRNALLNGDTKNYDWDSGYTCHQLG
	100-	<u> </u>			` <u> </u>	SGAIVVQLAQPYMIGSIRVLLWDCDDRSY
541	1891	A	4146	282	778	GTLGYPNGARGQPQDNFFAHQ\VSHHPPISAC

SEQ ID NO: of nucl- cotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine O=Cysteine, D=Aspartic Acid, E=Glutamic Acid, P=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first amino acid residue of peptide sequence	to last amino acid residue of peptide sequence	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
		٠				HAESENFAFWQDMKWKNKFWGKSLEIVPVG TVNVSLPRFGDHFEWNKVTSCIHNVLSGQRW IEHYGEVLIRNTQDSSCHCKITFCKAKYWSSN VHEVQGAVLSRSGRVLHRLFGKWHEGLYRG PTPGGQCIWKP
542	1892	A	4147	44	433	SVDAYVCNDIVFSYRTTITLLEGA*LTHRYVA QDPKQGQLRSLHLTCDSAPAGSQGTWSTSCR INHLIFRGGAQITFLATFDDSPKAVLGDRLLLT ANVSSENNIPRTSKTTFQLELSVKDAVYTVV SSH
543	1893	A	4153	678	11	TISYPQCLTQMYFLISFANVDTFLLPIMALDH YVAICSALQ*CSIITP/ELCQGLPVLA*AGSSLIS PVHTVIMSRLAFCSSAQISHFYRDAYLLMKIA CSHT*\NQHVFLGAVVLFLAPCALILVSYIRIA AAILRIPSPTRRKACSICSSHLSLVTLFYGTV LGICI*PPDSFSAQDAIATIMYTVVTSMLNPFTY SLMNKEVQEAVRRLFSRGSHSSWCW
544	1894	A	4158	3	538	LLYAQAGVQ*LNLSSLQPQPAGLKQSSHPSLP SSWDYRYSTPHPANFFVEMEFHHVAQAGLEL LGSGDLPTSTSHSAGITGV\SHHAPPRLISSEGS LLGHLLCLPMVFPLLCVFVLISSSLAGEEAAG LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE LKVGREGHVLPWQAHVVEF
545	1895	A	4160	1	412	HPLGLGLVPSEIFSPODKKAADGSILAPARGE DLEAGLKGSFMDGRLQASVSVFRIQRVGSAM QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGEPAPGWQVLAGYTYTQARYLRDASE ANVGQPLRPVDPR
546	1896	A	4174	1252	1190	FFQVFIFLFLIFFKTEFHSCCPGAVQWHDLDSL QPPPPRFKGFSCLSLPSSWDYRHAPAHPANFV FLVETGFLHV\GQ\ASLELPTSGDTPAS\ASQSA GITGVSHHA*PRASGRRCW
547	1897	À	4176	3029		AGPDGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSSVPLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA ALIVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASVPLCFRPAPGSLRQTPLPPQFHIPRGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPASPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSPLGPTSP MP*CFHRPSPPLP/LSSPTPA\LRPQAPQFPHHLP P*PPAPSPGCPLPPLAQPUPSPPSPHARSTLT PPLWPSLALLP*PLPPPPPVPSFSASLLCSLPAH

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		i .	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ľ	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
		ł	}		}	GTPASPGLGRSCLGKPQTLPWISFWPPSGRLA
						PGTWQPW/PVSPAPLSCLSAWDPWELPSPQPQ
]		l				VCSTAELPTSCLLSSPGP\PAFQPPRFGCL*GPP
	•	l				GPPGLPPLQSSLSFPPPPPPPVPQPPAPPALQWG
640	1000	<u> </u>	4100	222	244	LHLPGGRTK
548	1898	Α	4180	2369	844	RIHREEDFQFILKGIARLLSNPLLQTYLPNSTK
						KIQFHQELLVLFWKLCDFNKVGQPRGALQGD
						GEQLPQ*PGGRDSVRLRGVGQSCPSLELSPLG
		Į.				PSPHP*KFLFFVLKSSDVLDILVPILFFLNDAR
[						ADQSRVOLMHIGVFILLLLSGECNFGVRLNKP
l i						YSIRVPMDIPVFTGTHADLLIV\VFHKIITSGHQ
i !						RLQPLFDCLLTTVVNVSPYLKSLSMVTANKLL
				-		HLLEAFSTTWFLFSAAQNHHLVFFLLEVFNNI IQYQFDGNSNLVYAIIRKRSIFHQLANLPTDPP
1					·	TIHKALQRRRRTPEPLSRTGSOGGAPPWRAPA
l						PLPLQSQAPSRPVWWLLQALTS*PRSPRCOR
1						MAPCGPWNLSPSRAWRMAARLRGSPARHGG
				·		SSGDRP/HSSASGQWSPTPEWVLSWKSKLPLQ
1 1						TIMRLLQVLVPQVEKICIDKGLTDESEILRFLO
						HGTLVGLLPVPHPILIRKYQANSGTAMWFRT
						YMWGVIYLRNVDPPVWYDTDVKLFEIQRV
549	1899	Α	4191	858	321	LPWQRLGVLLSRGKMAVTGWLESLRTAQKT
1 1						ALLQDGRRKVHYLFPDGKEMAEEYDEKTSE
						LLVRKWRVKSALGAMGQWQLEVGDPAPLG
						AGNLGPELIKESNANPIFMRKDTKMSFQWRIR
				j		NLPYPKDVYSVSVDQKERCIIVRTTNKKYYK
550	1900	A	4192	1 -	1000	KFSIPDLDRHQLPLDDALLSFA\TPTAP
330	1900	^	4192	•	1980	IRHTGSDIAGVCGWLLLSGPCGVGLDLDSRLL
		ı				GASAMRRSEVLAEESIVCLQKALNHLREIWE
<u> </u>		- 1				LIGIPEDQRLQRTEVVKKHIKELLDMMIAEEE SLKERLIKSISVCQKELNTLCSELHVEPFOEEG
1 1	J	J	J			ETTILQLEKDLRTQVELMRKQKKERKQEVLKL
'						LQEQDQELC\EILCMPHYDIDSASVPSLEELNO
					]	FRQHVTTLRETKASRREEF/VSSIKRQIILCME
]	:					ELDHTPDTSFERDVVCEDEDAFCLSLENIATYL
] ]		I	ŀ			QKLLRQ\LEMQKSQNEAVCEG\LRTOI\RELW
[	}	- 1	ľ	1	ľ	DRLQIPEEEREAVATIMSGSKAKVRK\ALQ\LE
	İ	1		]	0	VDRLEELEKCKTMKKVIEAIRVELVQYWDQC
		l	ł	ļ		FYSQEQRQAFAPFCAEDYTESLLQLHDAEIVR
						LKNYYEVHKELFEGVQKWEETWRLFLEFER
1 1		ł	l	I	ł	KASDPNRFTNRGGNLLKEEKQRAKLQKMLP
	1	1	*		1	KLEEELKARIELWEQEHSKAFMVNGQKFME
{		1	ŀ	, [	1	YVAEQWEMHRLEKERAKQERQLKNKKQTET
		ļ	l		. [	EMLYGSAPRTPSKRRGLAPNTPGKARKLNTT
1 1	1	i	ļ	. 1	j	TMSNATANSSIRPIFGGTVYHSPVSRLPPSGSK
	·	1	- 1	**		PVAASTCSGKKTPRTGRHGANKENLELNGSI
]		j		<u> </u>		LSGGYPGSAPLQRNFSINSVASTYSEFADPSLS
551	1901	A	4194	3	1008	DSSTVGLQRELSKASKSDATSGILNSTNIQS
	1701	^ ]	7177	-	1000	AWHEGLVSSPAIGAYLSASYGDSLVVLVATV
		- 1		ĺ	-	VALLDICFILVAVPESLPEKMRPVSWGAQISW KQADPFASLKKVGKDSTVLLVCITVCLSYLPE
	- 1	- 1			· [	AG\QYSSFF\LYLR\QVIGFG\TVKIAAFIAMVGI
	- 1	1		. [	ł	LSIVAQTAFLSILMRSLGNKNTVLLGLGFOML
J J	J	1	- 1	1	ļ	QLAWYGFGSQAWMMWAAGTVAAMSSITFP
	.	. [	1	i	į	AISALVSRNAESDQQGVAQGIITGIRGLCNGL
	j	- 1	I	l	i	GPALYGFIFYMFHVELTELGPKLNSNNVPLQ
l. I	1	- 1		ŀ		GAVIPGPPFLFGACIVLMSFLVALFIPEYSKAS
.		ĺ			l	GVQKHSNSSSGSLTNTPERGSDEDIEPLLQDS
[ [	- 1	ļ	1		ł	SIWELSSFEEPGNOCTEL

1902 A   4197 2   14302   14	SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
AVKEGSLHARDPVGGTTEFLEVPI IKI EVTI I I	552	1902	A	4197			ARPPPAPGSRQQKQKAAPGAAAAAELRGAR EPAPARRGTMADGGEGEDEIQFLRTDDEVV LQCTATIHKEQQKLCLAAEGFGNRLCFLESTS NSKNVPPDLSICTFVLEQSLSVRALQEMLANT VEKSEGQVDVEKWKFMMKTAQGGGHRITL YGHAILLRHSYSGMYLCCLSTSRSSTDKLAFD VGLQEDTTGEACWWTIHPASKQRSEGEKVR VGDDLILVSVSSERYLHLSYGNGSLHVDAAF QQTLWSVAPISSGSEAAQGYLIGGDVLRLLH GHMDECLTVPSGEHGEEQRRTVHYEQGAVS VHARSLWRLETLRVAWSGSHIRWGQPFFLLR HVTTGKYLSLMEDKNLLLMDKEKADVKSTA FTFRSSKEKLDVGVRKEVDGMGTSEIKYGDS VCYIQHVDTGLWLTYQSVDVKSVRMGSIQR KAMHHEGHMDDGISLSRSQHEESRTARVIRS TVFLFNRFIRGLDALSKKAKASTVDLPIESVSL SLQDLIGYFHPPDEHLEHEDKQNRLRALKNR QNLFQEEGMINLVLECIDRLHVYSSAAHFAD VAGREAGESWKSILNSLYELLAALIRGNRKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIIKEGHIKSIISLLDKHGRNHKVLD VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEEWGGNGVGDDLFSVGFDGLHLWSG CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMFENINIDGLFFPVVSFSAGIKV RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGPTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENIHELWYMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLALGGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLIGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTIMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGYAQVGRMNFGKDVSTLKYPTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFHMQTSTYYYSVRIFFGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPCCPPRLHVQFLSHVLWSRMPNOFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEGELLKFHYHTILRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSTILFPDENKKHGLPGIGSTSLRPRMQF SSPSFVSISNECYQYSPEPFLDILKSKTIOM.TR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	De A mortio A 14 P. Charmin A 14
		nou				D-Aspartic Acid, E-Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	l			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	l			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ì					acquaice	
1	<b>[</b>			peptide		/=possible nucleotide deletion, \=possible
<u> </u>				sequence		nucleotide insertion
1	•					MGIFHNEDLKHILQLIEPSVFKEAATPEEESDT
1			1			LEKELSVDDAKLQGAGEEEAKGGKRPKEGLL
	t :					QMKLPEPVKLQMCLLLQYLCDCQVRHRIEAI
1			]			VAFSDDFVAKLQDNQRFRYNEVMQALNMSA
		•				ALTARKTKEFRSPPQEQINMLLNFKDDKSECP
						COCCIDENT A DESCRIPTION OF THE COLUMN PROPERTY OF THE COLUMN PROPERT
			1		1	CPEEIRDQLLDFHEDLMTHCGIELDEDGSLDG
						NSDLTIRGRLLSLVEKVTYLKKKQAEKPVES
						DSKKSSTLQQLISETMVRWAQESVIEDPELVR
1	1 1					AMFVLLHRQYDGIGGLVRALPKTYTINGVSV
1						EDTINLLASLGQIRSLLSVRMGKEEEKLMIRG
						LGDIMNNKVFYQHPNLMRALGMHETVMEV
ļ			1			MVNVLGGGESKEITFPKMVANCCRFLCYFCR
1	l İ					ISRQNQKAMFDHLSYLLENSSVGLASPAMRG
1	J J					STPLDVAAASVMDNNELALALREPDLEKVVR
1	]	- 1		·		VI ACCOLOGOOM VOICEMENTATION TO THE
			.	J	ļ	YLAGCGLQSCQMLVSKGYPDIGWNPVEGER
			ļ	J		YLDFLRFAVFCNGESVEENANVVVRLLIRRPE
1		- 1		·		CFGPALRGEGGNGLLAAMEEAIKIAEDPSRD
		]			ļ	GPSPNSGSSKTLDTEEEEDDTIHMGNAIMTFY
		l		-		SALIDLLGRCAPEMHLIHAGKGEAIRIRSILRS
J		ì				LIPLGDLVGVISIAFQMPTIAKDGNVVEPDMS
		ſ	ſ		•	AGFCPDHKAAMVLFLDRVYGIEVQDFLLHLL
1 1		- 1	- 1	i		EVGFLPDLRAAASLDTAALSATDMALALNRY
1	1	į	l			LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT
]	1	- 1	ł	ł		VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS
1		- 1	ŀ	į	i	MMQHLLRRLVFDVPLLNEHAKMPLKLLTNH
1 1		- 1	- 1	-		
			1			YERCWKYYCLPGGWGNFGAASEELHLSRK
1 1			- 1	]		LFWGIFDALSQKKYEQELFKLALPCLSAVAG
1			i	i	•	ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ
1	1					PVDTSNITTPEKLEYFINKYAEHSHDKWSMDK
1 1			J	j	J	LANGWIYGETYSDSSKVQPLMKPYKLLSEKE
				ı		KEIYRWPIKESLKTMLARTMRTERTREGDSM
i I				Į.		ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS
i 1		ļ	. [	ŀ		NVTLSRDLHAMAEMMAENYHNIWAKKKKM
i I		l			İ	ELESKGGGNHPLLVPYDTLTAKEKAKDREKA
1 1				1		QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA
[ . [	1	ŀ	4	ı	1	YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF
1	ł	l	l	j		PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA
	1	- 1	j	l		SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH
j	ĺ	- 1	i	l		
	l	- 1	l	l		RISLFGNDATSIVNCLHILGQTLDARTVMKTG
	ľ		1	l	ł	LESVKSALKAFLDNAAEDLEKTMENLKQGQF
j	ŀ	ı		l	i	THTRNQPKGVTQIINYTTVALLPMLSSLFEHI
	1	- 1	1	ľ	ı	GQHQFGEDLILEDVQVSCYRILTSLYALGTSK
	I	ł	ŀ		i	SIYVERQRSALGECLAAFAGAFPVAFLETHLD
1	J	ŀ	i			KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP
	· 1	1			ŀ	SLEKLMEEIVELAESGIRYTQMPHVMEVILPM
				· [	l	LCSYMSRWWEHGPENNPERAEMCCTALNSE
1 1	ļ	Į	j	J	J	HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF
	ļ	- 1				SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV
j l	ŀ	-	1			SEEDHLKAEARGDMSEAELLILDEFTTLARDL
1. 1	[	l		.		
	1	ł				YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR
	I	1		İ		MVAEVFIYWSKSHNFKREEQNFVVQNEINN
1 1	ſ	ł	- }	ŀ		MSFLITDTKSKMSKAAVSDQERKKMKRKGD
1 ł	i	J			1	RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA
{	ļ	1		1		LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP
j. j		ł	. 1	1		AIRWOMALYKOLPNRTDDTSDPEKTVERVL
ļ <b>1</b>	ŀ		ļ	Ī		DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK
į J	J	l	1			KAVWHKLLSKQRKRAVVACFRMAPLYNLPR
ş l				1	[	HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA
}- I	1			l		KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT
	1			l		EVCVI FEDEL VALAVADDA A VOCUMENTALI
				L	l	EKCKLEEDFLYMAYADIMAKSCHDEEDDDG

SEQ NO:	of NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	1	ļ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotid	1	1	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	<sup>E</sup>	1.	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	ł	ł	]	amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ĺ	İ	i	ł	peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
1	1		ľ	sequence		nucleotide insertion
<b>—</b>		<del> </del>	<del> </del>	Sequence		EEEVKSFEEKEMEKQKLLYQQARLHDRGAA
Ì	<b>]</b> .					EMVLQTISASKGETGPMVAATLKLGIAILNGG
1	1	Í	1		j	NSTVQQKMLDYLKEKKDVGFFOSLAGLMOS
1		1	ì			CSVLDLNAFERQNKAEGLGMVTEEGSGEKV
1	1		Ì	ļ		LQDDEFTCDLFRFLQLLCEGHNSDFQNYLRT
1	1 .	1		ľ		QTGNNTTVNIIISTVDYLLRVQESISDFYWYY
1	1	ł	)			SGKDVIDEQGQRNFSKAIQVAKQVFNTLTEYI
		1	1		1	QGPCTGNQQSLAHSRLWDAVVGFLHVFAHM
}	ľ	ľ	ì	i		QMKLSQDSSQIELLKELMDLQKDMVVMLLS
1	- 1	ĺ	ļ	ļ		MLEGNVVNGTIGKQMVDMLVESSNNVEMIL
1	*	1		1		KFFDMFLKLKDLTSSDTFKEYDPDGKGVISK
						RDFHKAMESHKHYTQSETEFLLSCAETDENE TLDYEEFVKRFHEPAKDIGFNVAVLLTNLSEH
ł	0.0	1	1	l		MPNDTRLQTFLELAESVLNYFQPFLGRIEIMG
			1			SAKRIERVYFEISESSRTQWEKPQVKESKRQFI
1				ĺ		FDVVNEGGEKEKMELFVNFCEDTIFEMQLAA
1	ŀ	1				QISESDLNERSANKEESEKERPEEQGPRMAFF
		1				SILTVRSALFALRYNILTLMRMLSLKSLKKQM
-	- 1	ĺ	<b>-</b>	ĺ		KKVKKMTVKDMVTAFFSSYWSIFMTLLHFV
1	i					ASVFRGFFRIICSLLLGGSLVEGAKKIKVAELL
		1				ANMPDPTQDEVRGDGEEGERKPLEAALPSED
1		1	<b>1</b>			LTDLKELTEESDLLSDIFGLDLKREGGQYKLIP
l	1	1	]			HNPNAGLSDLMSNPVPMPEVQEKFQEQKAK
i .						EEEKEEKEETKSEPEKAEGEDGEKEEKAKED KGKQKLRQLHTHRYGEPEVPESAFWKKIIAY
						QQKLLNYFARNFYNMRMLALFVAFAINFILL
İ	•	ł				FYKVSTSSVVEGKELPTRSSSENAKVTSLDSS
i .		1	]			SHRIIAVHYVLEESSGYMEPTVRILPILHTVISF
(	l					FCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI
l			ŀ			TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF
ļ	-			•		VKRKVMDKYGEFYGRDRISELLGMDKAALD
( .		i	i 1			FSDAREKKKPKKDSSLSAVLNSIDVKYQMW
553	1903	A	4199	31	767	KLGVVFTDNSFLYLAWYMT
333	1903	^	4199	31	/6/	LPELNGRGAGLRRAEPSERGGGAERTQQVAA
(	1	Ì	1 .			LPLSHGHSHGGGGCRCAAER/VGAARGSAAC AYGLYLRIDKGRLOCLNESREGSGRGVFKPW
1		,	] ]	l		ERAD\DRSKFVESDADEELLFNIPFTG\HVKLK
						GIIMGEDDDSHPSEMRLYKNIPQMSFDDTER
		1		·		EPDQTFSLNRDLTGELEYATKISRFSNVYHLSI
ļ		] .		, <b>j</b>	• ]	HISKNFGADTTKVFYIGLRGEWTELRRHEVTI
<u> </u>						CNYEASANPADHRVHQVTPQTHFIS
554	1904	A	4200	1	961	GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSL
l		1			ļ	EICIKACKNLAYGEEKKKKCNPYVKTYLLPD
1		1				RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA
			1 1			QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT
1						WDFEDSTTQSFRWHPLRAKADKYEDSVPQS
[		1				NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL
<b>!</b>	1	1				HGQLCLVVLGAKNLPVRPDGTLNSPVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV
	1	]				TPAQLRQSSLELTVWDQALFGMNDRLLGGT\
	1		[	1		RLGSKGDTAVGGDACSQSKLQWQKVLSSPN
				į		LWTDMTLVLH
555	1905	A	4211	331	2419	KENKKARNLRMNQSRSRSDGGSEETLPQDH
	- [			[		NHHENERRWQQERLHREEAYYQFINELNDE
	1	-			ļ	DYRLMRDHNLLGTPGEITSEELQQRLDGVKE
]		1	1 1			QLASQPDLRDGTNYRDSEVPRESSHEDSLLE
	J	1 1				
I	- 1	1		ſ	İ	WLNTFRRTGNATRSGQNGNQTWRAVSRTNP
						NNGEFRESLEIHVNHENRGFEIHGEDYTDIPLS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  GGAAGIPRANASRTNFSSHTNQSGGSELRQRE GQRFGAAHVWENGARSNVIVRNTNQRLEPI RLRSTSNSRSRSPIQRQSGTVYHNSQRESRPV QQTTRRSVRRGRTRVFLEQDRERERGTAY TPFSNSRLVSRITVEEGEESSRSSTAVRHPTIT LDLQVR\RIRPGENRDRDSIANRTRSRVGLAE NTVTIESNSGGFRRTISRLERSGIRTYVSTITVP LRRISENELVEPSSVALRSILRQIMTGFGELSSL MEADSESELQRNGQHLPDMHSELSNLGTDN
556	1906	A	4212	3	462	NRSQHREGSSQDRQAQGDSTEMHGENETTQP HTRNSDSRGGRQLRNPNNLVETGTLPILRLAH FFLLNESDDDDRIRGLTKEQIDNLSTRHYEHN SIDSELGKICSVCISDYVTGNKLRQLPCMHEF HIHCIDRWLSENCTCPICRQPVLGSNIANNG LQRQRQHPAAAPAVPVRCFTFCFTDIVIMPKR
						KSPENTEGKOGSKVTKQEPTRRSARLSAKPA PPKPEPKPRKTSAKKEPGAKISRGAKGKKEEK QEAGKEGTAPSENGETKAEEIHISRSTVNVST SRGTPPSTLSVKGQIETVRVKGTEN
557	1907	A	4213	774	507	ARRFSCLTLQTSWGHRH\GPPRP\ANFVFLVET GFLHIGQAGHKLPTSGDPPASASQSARITGMS HRTWFLASFLIDSCKNFTVVKIMVTI
558	1908	A	4225	3	1253	HRTWFLASFLIDSCKNFIVYKIMYTL  TYRHAEREHPETSSATKVSYDYRHKRPKLLD GDQDFSDGRTQKYCKEEDRKYSFQKGPLNRE LDCFNTGRGRETQDGQVKEPFKPSKKDSIAC TYSNKNDVDLRSSNDKWKEKKKEGDCRKE SNSSSNQLDKSQKLPDVKPSPINLRKKSLTVK VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT IIHQVKANYFPSPGITLHERFSVKMADIHKADV NEIPLNSDPEIHRRIDMSLAELQSKQAVIYESE QTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAERDDQNSSPSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFTTHKFRERLMRK KKVP
	1909	Α	4235	1	323	KFSIPFFLRWSFTLVYPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS
560	1910	A	4246		1569	TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LL'NDFLRTGLLICGNGK\FHKHLQDLFAPLVV RYMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KLCSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD
561	1911	A	4257	1300	654	SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI FPDLFKRAAERLQYVFGYKLVELEPKSNTYIL

WO 01/57188

NO: of nucleotide cotide sequence peptide sequence uncleotide lusts a nucleotide cotide sequence uncleotide sequence uncleotide sequence peptide sequence uncleotide lusts an incleotide sequence uncleotide sequence peptide seque	steine	Amino acid sequence (A=Alanine C=Cystei	Predicted end	Predicted	SEQ	Met	SEQ ID	SEQ ID
Ducitod cotide   Sequence   USSN   Gostion	u,	D=Aspartic Acid. E=Ghutamic Acid.						
cotide sequence   USSN   tocation   corresponding to last amino acid residue of peptide   residue of peptide   sequence	ne.					ļ .	peptide	nucl-
Sequence   Uence   09/496   corresponding to first amino acid residue of peptide residue of peptide sequence   Peptide sequen	,			docation	USSN	1	seq-	cotide
uence   914   ng to first amino acid residue of peptide residue of peptide sequence   T=Threonine, V=Valine, W=*Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, v=possible nucleotide insertion   NTLEPVEEDAEMRGDQGTPTTGLIMF   PMKGMTIKETEAWDPLIALIGVYPTKK   DPKKLITEDFVRQRYLEYRRIPHTDPVI   WGPRTINLETSKMKVILEYRAKVHNQDI   PAQYCEALADEENRARPQPSGPAPSS   WGPRTINLETSKMKVILEYRAKVHNQDI   PAQYCEALADEENRARPQPSGPAPSS   WHARLAGESTLERPILALIGVYPTKK   DPKKLITEDFVRQRYLEYRRIPHTDPVI   WGPRTINLETSKMKVILEYRAKVHNQDI   PAQYCEALADEENRARPQPSGPAPSS   WGPRTINLETSKMKVILEYRAKVHNQDI   PAQYCEALADEENRARPQPSGPAPSS   WGPRTINLETSKMKVILEYRAKVHNQDI   PAQYCEALADEENRARPQPSGPAPSS   WGPRTINLETSKMCKFLSRGFCCEPHYLFLA   SLSTSAISFAEVQVQAPPVVAATPSPTA   ASGETADVVQTAAEQSFAELGLGSYTF   QNLEEPMIVDLGLEPWGAAACTVPA   PLIVTGQRRAARTHNHLPEQKTSSRRE   GDHEYYKASSEMALYQKKHGIKLYKI   TQAPIFISFIFIALRMANI,PYPSLGTGGL   QDLTVSDPTYILPLAVTATMWAVLELG   VQSSDLQWMRNVRMMPLITI.PITMIFF   PMYWLSSNI,FSLQVQKCRIPAVRYUL   VVHDLDKLPPREGFLESFKKGWKNAE!   LREREQRMRNQLELAARGPLRQTTTIM   PGKONPPNIPSSSSSSSSNSSNI,FSLQVGKRAPER   VFRAYSHI   VFRAY	10	M=Methionine N=Asparagine P=Proline			09/496		uence	seq-
amino acid residue of peptide sequence Peptide sequence Peptide sequence Peptide sequence Peptide sequence Peptide sequence Peptide sequence Peptide sequence Peptide sequence Peptide sequence Peptide insertion Interpretation Peptide insertion Interpretation Peptide insertion Interpretation Peptide insertion Interpretation Peptide insertion Period Peptide Peptide Peptide Interpretation Pe	,					ì	1	
residue of peptide sequence peptide sequence peptide nucleotide deletion, t-possible nucleotide insertion interpretation in the period in thep		T=Threonine V=Valine W=Tryntonhan						
peptide sequence    Possible nucleotide deletion, '-possible nucleotide insertion   InflierveEdDAEMRGDQGTPTTGLIMT   InflierveEdDAEMRGDQGTPTTGLIMT   FMKGNTIKETEAWDPLLALGVPYTKK   DPKKLITEDPVRQRYLERPHTDPPVQ   WOPRITNLETSKMKVI.KFVAKVHNQDI   PAQYCEALADEENRARQPGSGAPSS					1	l		
	4 <b>)</b>	=nossible nucleotide deletion \-nossible	Sequence			l		
INTLEPVEEDAEMRGDQGTPTTGLLMT	5					1		
PMKGNTIKETEAWDFLLALIGVYPTIKE	MWI CI			Бодаблос	<del> </del>			
DPKKLITEDFVRQRYLEYRRIPHTIDPUT   WGPRTNLETSKMKVLKFVAKVHNQDI   PAQYCEALADEENRARPQPSGPAPS    S62								
WGPRTINLETSKMKVLKFVAKVHINQDE   PAQYCEALADEENRARPQPSGPAPSS								
PAQYCEALADEENRARPPSGPAPSS					1		•	
1912   A   4260   1   1498   MYTWLYRFLPTSNMAKKRSLLPPDLI   WLHARLQKCFLSRGCGSYCAGAKASP   WAGKPLTTRILIFPAAPCCCRPHYLFLA   SLSTSAISFAEVQVQAPPVVAATPSPTA   ASGETADVVQTAAEQSFAELGLGSYTT   QNLLEFMHVDLGLPWWGAIAACTVFA   PLIVTGQREARHNILPEKQFKSSRRE   GDHIEYYKASSEMALYQKKHGIKLYKI   TQAPPTISFTIALREMANLPVPSLQTGGL   QDLTYSDPYLIPLAVTATMWAVLELG   QVGSDLQWMRNVIRMMPLITLPITMHF   FMYWLSSNLFSLVQVSCLRIPAVRTVLI   VVHDLDKLPPREGFLSFFKRGWKNAE   LREREQRMRNQLELAARGPLRQTFTH   PGKDMPPNIPSSSSSSSKYKSKYPWHDT   THOTRETLSSOFSNLFILPLSSSATMPS   HRSPNGGLFRGSPVK/TPPPMSFQPVC   PRGSGNPPHGTSILTAPPALLPHPPTHPT   LIQENNTINHHSHTHTYTETLSFFLYK   DRMEWGKSY   PREMEWGKSY   PREMEWGKSY   FIFLYVCLLSJCQVQKQYQKWFREIVKS   ETYTLSKMGPDSKPSGDVFPRTSE   FIFLYSCLSJQQVQKQYQKWFREIVKS   ETYTLSKMGPDSKPSGDVFPRTSE   FIFLYCLLSJCQVQKQYQKWFREIVKS   ETYTLSKMGPDSKPSGDVFPRTSE   PSALLAPTKPRALGTLRLYECSPELGT   PAWLLMLCQAPRPQDPDPRLTQPEKSL   GYRYLTILFTCSTPWABSSTKFRHIMY   GLTSFGEKVEFLNIGMMDLSYASD   VLEVSQAPVIFSHSAARAVCDNLLNVPI   LIKKNGGIVMVTLSMQLQCLLANV   DHPDHRAVIGSEFIGIGGRYDGTGFFF   DVSTYPVLIEELLSRSWSSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   PHLGASSWYSEE								
WLHARLQKCFLSRGGSYCAGAKASP    MAMGLMCGRELILLQSGRRVHSVV    WLGKPLTTRLIFPAAPCCGRPHYLFLA    SLSTSAISFAEVQVQAPPVVAATPSPTA    ASGETADVVQTAAEQSFAELGI.GSYTF    QNLLEPMHYDLGLPWWGAIAACTVFA    ASGETADVVQTAAEQSFAELGI.GSYTF    QNLLEPMHYVASSEMALYQKKHGIKLYKI    TQAPITISFIALREMANLPYPSL.QTGGL    QDLTYSDPIYILPLAYTATMWAVLELG    QDLTYSDPIYILPLAYTATMWAVLELG    QVSSDLQWMRNVIRMMPLITLPITIMFF    FMYWLSSNLFSLVQVSCLRIPAVRTVIL    VVHDLDKLPPREGFLESFKKGWKNAE    LRERGMRNQLELAARGPLRQTIFTH    PGKONPPHIPSSISSSSSKKKSKYPWHDI    THOTRRTLSGPSNNLPPLSSSATMPSI    HRSPNGGLFRQSPVK/TPPIPMSFQPVOF    PRGSGNPPHGTSILTAPPALLPHPFTHFILQENNNTHIHSHTHTYTETLSFFLYIC   DRMEWGKSVF    FMSCHALQVGPAAQVMAYLFTINS    S64			1400	1	4260	<del></del>	1012	562
MAMGLMCGRRELILLLQSGRRVHSVA   WLGKPLTTRLLFPAAPCCCRPHYLFLA   SLSTSAISFAEVQVQAPPVVAATPSPTA   ASGETADVVQTAAEQSFAELGI.GSYTF   QNILLEFMHYDLGLPWWGAIAACTVFA   PLIVTGQREAARI:NHLPEIQKFSSRIRE   GDHEYYKASSEMALYQKKHGIKLYKI   TQAPFIISFIAREMANLPYPSL.QTGGL   QDLTVSDPIYILPLAVTATMWAVLELG   VQSSDLQWMRNVIRMPLITLPTIMHF   PMYWLSSNLFSLVQVSCLRPAVRTVLL   VVHDLDKLPPREGFLESFKKGWKNAE!   LREREQRMRNQLELAARGPLRQTFTHH   PGKONPPNIPSSISSSSSKPKSKYPWHDT   VVHDLDKLPPREGFLESFKKGWKNAE!   LREREQRMRNQLELAARGPLRQTFTHH   PGKONPPNIPSSISSSSSKPKYWHDT   LQENNINHHTHSHTHTYTETLSFFLYIC   PRGSGNPFHGTSIL TAPPALLPHPPTHPT   LQENNINHHTHSHTHTYTETLSFFLYIC   DRMEWGKSVF   LLKRKLSSLNSEVSTIQNTRMLAFKATA   GCTWCLGLJQVQKQYQKWFREIVKS   ETYTLSKMGPDSKPSEGDVFPRTSE   S65   1915   A 4288   83   406   RNSRPLWCSPASOPRQAPVGSCCOPE   PSALLAPTKPRALGTLRLYECSPELGT   PAWLLMLCQAPRQDPDPRLTQPEKSL   GQTGASRTPRT   GURYLTLIFTCSTPWAESTKFRHHMY   GLTSFGEKVEELNRLGMMDILSYASD   VLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCALLANV   DHTDHIRAVIGSEFIGIGGRYDLLANV   DHTDHIRAVIGSEFIGIGGGRYDLANV   DHTDHIRAVIGSEFIGIGGGRYDLANV   DHTDHIRAVIGSEFIGIGGGRYDLANV   DHTDHIRAVIGSEFIGIGGGRYDLANV   DHTDHIRAVIGSEFIGIGGGRYDLANV   DHTDHIRAVIGSEFIGIGGGRYDLANV   DHTDH			1498	1	4200		1912	302
WIGKPLTTRLLFPAAPCCCRPHYLFLA SLSTSAISFAEVQVQAPPVVAATPSPTA ASGETĄDVVQTAABQSFAELGLGSYTF QNILBEMHYDLGLPWGAIAACTVFA PLIVTGQREAARIHNHLPEIQKFSSRIRE GDHIEYYKASSEMALYQKKHGIKLYKI TQAPIFISFFIALREMANLPVPSLQTGGL QDLTVSDPIYILPLAVTATIMAVLELG QDLTVSDPIYILPLAVTATIMAVLELG QDLTVSDPIYILPLAVTATIMAVLELG QDLTVSDPIYILPLAVTATIMAVLELG VQSSDLQ WMRNVIRMMPLITIPITIMIF FMYWLSSNLFSLVQVSCLRIPAVRTVLI VVHDLDKLPPREGFLESFKKGWKNAEI LREREQRMRNQLELAARGPLRQTFTIH PGKDNPPNIPSSISSSSKPKSK YPWHDT HGGRAPTQTLEPTNEYQNTQLSVSYLI THGTERTLSSGPSNNLPLPLSSSATMPSI HRSPNGGLFRQSPVK/TPPIPMSPQPVPC PRGSGNPPHGTSILTAPPALLPHPPTHPT LIQENNNTNHTHSHTHTYTETLSFPLYIC DRMEWGKSVF PRGSGNPPHGTSILTAPPALLPHPPTHPT LIQENNNTNHTHSHTHTYTETLSFPLYIC DRMEWGKSVF FILVYCLLSIQQVQKQYQK WFREIVKS ETYTLSKMGPDSKPSEGDVFPRTSE ETYTLSKMGPDSKPSEGDVFPRTSE 565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPI PPSALLAPTKPRALGTLRLYECSPELGT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASKTPRT GVTSLTHTCTSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMDLSYASD VLEVSQAPVFSHSAARAVCDNLLNVIL LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LLKKNGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKRGGIVMVTLSMGVLQCNLLANVI LHKGGIVMVTLSMGVLQCNLLANVI LHKGGIVMVTLSMGVLQCNLLANVI LHKGGIVMVTLSMGVLQCNLLANVI LHKGGIVMVTLSMGVLGCNLLANVI LHKGGIVMVTLSMGVLGCNLLANVI LHKGGIVMVTLSMGVLGCNLLANVI LHKGGIVMVTLSMGVLGCNLLANVI LHKGGIVMVTLSMG	SPLPGK	WLHARLQKCFLSKGCGSYCAGAKASPI			!			
SLSTSAISFAEVQVQAPPVVAATPSPTA ASGETADVVQTAAEQSFAELGLGSYTT QNLLERMHVDLGLPWWGAIAACTVFA PLIVTGQREAARIHNHLPEIQKFSSRIRE GDHIEYYKASSEMALYQKKHGIKLYKI TQAPFISFFIALREMANIPVPSLQTGGL QDLTVSDPIYILPLAVTATMWAVLELG VQSSDLQWMRNVIRMMPLITLPITMHF FMYYLLSSIN,SLVQVSCLIPPAVRTVLI VVHDLDKLPPREGFLESFKKGWKNAEI LREREQRMRNQLELAARGPLRQTFTTH PGKDNPPNIPSSISSSSKPKSK YPWHDT THOTTRILSSGPSNNLPLPLSSSATMPSI HRSPNGGLFRQSPVK/TPPIPMSFQVPVG PRRGSCNPPHGTSILTAPPALLPHPTHFT LIQENNITHHTHSTHTYTETLSFFLYIG DRMEWGKSVF  1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGILQVGPAAQVMAYLFTINS FIFLVYCLLSIQQVQKQYQKWFREIVKS STYYLLSSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQFRQAPVSQCCCPL PPSALLAPTKPRALGTLRLYECSPELGT PAWLLML_CQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GYRYLTLTTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVFFSHSAARAVCDNLLNVPI LLKKNGGIVMYTLSMGVLQCNLLANVI DHFDHIRAVIGSEFIGIGGNYDGTGFFF DVSTYPVLIEELLSSW SEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR	VAGPSC	MAMGLMCGRRELLRLLQSGRRVHSVA	}		1			
ASGETADVVOTĀAEQSFAELGLGSYTTE QNLLEFMHVDLGLPWWGAIAACTVFA PLIVTGQREAARHNHLPEIQKFSSRIRE GDHIEYYKASSEMALYQKKHGIKLYKI TQAPFISFĪALREMANLPVPSLQTGGL QDLTVSDPIYILPLAVTATMWAVLELG VQSSDLQWMRNVIRMMPLITLPITMHE PMYWLSSNLFSLVQVSCLRFAVRTVIL VVHDLDKLPPREGFLESFKKGWKNAEI LÆREQRMRNQLELAARGPLRQTFTHN PGKDNPPNIPSSSSSSKKKYPWHDI  563 1913 A 4265 623 116 MGGLAPTQTLEPTREYQNTQLSVSYLI THGTRRTLSSGPSNNLPLPLSSSATMPSI HRSPNGGLFRQSPVK/TPPIPMSPQPVPG PRGSGRPPGTSIL TAPPALLPHPPTHFT LIQENNNTNHTHSHTHTYTETLSFFLYIG DRMEWGKSVF  564 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGLLQVGPAAQVMAYLFTINS FIFLVYCLLSIQQVQKQYQK WFREIVKS ETYTLSSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQFRQAPVSQSCCCPI PPSALLAPTKPRALGTLRLYECSPELGT PAVLLMLCQAPRPQDPPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVYBALAPTKPRALGTLRLYECSPELGT PAVLLMLCQAPRPQDPPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVYBALTTTTCSTPWAESSTKFRHHMY GLTSFGEKVVEELINRLGMMDLSYASD VLEVSQAPVIFSHSAARAVCDNILLNVPI LLKKNGGIVMVTLSMGVLQCNLLANV DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS								
QNLLEFMHYDLGLPWWGAIAACTYPA PLIVTGQREAARIHNHLPEIQKFSSRIRE GDHEYYKASSEMAL YQKKHGIKLYKI TQAPIFISFFIALREMANLPVPSLQTGGL QDLTVSDPYJHLPAVTATMWAVLELG VQSSDLQWMRNVIRMPLITLPITMIF FMYWLSSNLFSLVQVSCLRIPAVRTVLI VVHDLDKLPPREGFLESFKKGWKNAEI LREREGMRNQLELAARGPLRGTFTHN PGKDNPPNIPSSSSSSSKKKSYPWHIDT THORRTLSSGPSNNLPPLSSSATMPSI HRSPNGGLFRQSPVK/TPPPMSFQPVPG PRGSGNPPHGTSIL TAPPALLPHPPTHPT LIQENNNTNHHSHTHTYTETLSFFLYIG PRMEWGKSVF  564 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGLLQVQPAAQWMAYLFTINS FIFLYCLSIQQVQKQYQKWFREIVKS ETYTLSSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPI PPSALLAPTKPRALGTLRLYECSPELCT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GYRYLTLTFTCSTPWAESSTKFRHMMY GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNLLNVPI LLKKNGGIVMVTLSMGVLQCHLLANV DHFDHIRAVIGSEFIGIGGNYDGTGFFC DVSTYPYLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS	<b>TAVPEV</b>	SLSTSAISFAEVQVQAPPVVAATPSPTA						
PLIVTGQREAARIHNHLPEIQKFSSRIRE GDHLEYYKASSEMALYQKKHGIKLYKI TQAPIFISFIALREMANLPVPSLQTGGL QDLTVSDPIYILPLAVTATMWAVLELG VQSSDLQWMRNVIRMMPLITLPTIMHE FMYWLSSNLFSLVQVSCLRIPAVRTVLI VVHDLDKLPPREGFLESFKKGWKNAEI LREREQRMRNQLELAARGPLRQTFTHN PCKDNPPNIPSSSSSSKFKSYPWHIDT THGTRRTLSSGPSNNLPLPLSSSATMPSI HRSPNGGLFRQSPVK/TPPIPMSFQPVPC PRGSGNPPHGTSILTAPPALLPHPPTHPT LIQENNTNHTHSHTHTYTETLSFFLYIG DRMEWGKSVF  564 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLOLLQVQPAAQVMAYLFTIINS FIFLVYCLLS\QQVQKQYQKWFREIVKS ETYTLSSKMGPDSKPSEGDVFPRTSE 565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAVGNSCCCCPT PPSALLAPTKPRALGTLRLYECSPELCT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSQKLACLIGVEGGHSLDSSLSVLRS GYRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELINLGMMDLSYASD VLEVSQAPVIFSHSAARAVCDNLLNVPI LLKKNGGIVMYTLSMGVJQCALLANV DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPYLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS	TPVGLI	ASUETADYVQTAAEQSFAELGLGSYTP		1				
GDHIEYYKASSEMAL YQKKHGIKL YKI TQAPIFISFIALREMANI PVPSLQTGGL QDLTVSDPIYILPLAVTATIMWAVLELG VQSSDLQ WMRNVIRMMPLITLPITMIFI PMYWLSSNLYSLVQVSCLRIPAVRTVLI VVHDLDKL PPREGFLESFKKGWKNAE LREREQRMRNQLELAARGPLRQTFTHN PGKDNPPNIPSSSSSSKYKSK YPWHDT PGKDNPPNIPSSSSSSSKYKSK YPWHDT THOTRITLSSGPSNNLPLPLSSSATMPSS HRSPNGGLFRQSPVK/TPPIPMSFQPVPG PRGSGNPPHGTSIL TAPPALLPHPPTHIPT LQENNINHIHHSHTHYTETLSFFLYIG DRMEWGKSVF  1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGLQVQPAQVWAYLFTIINS FIFLVYCLLSQQVQKQYQK WFREIVKS ETYTLSSKMGPDSKPSEGDVPPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPI PSALLAPTKPRALGTILI YECSPELGT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLITFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMDLSYASD VLEVSQAPVIFSHSAARAVCDNILINVPI LLKKNGGIVMYTLSMGVLQCNILANVI LLKKNGGIVMYTLSMGVLQCNILANVI DHFDHIRAVIGSEFIGIGGNYDGTGRFPQ DVSTYPYLIEELLSRSWSEELQGVLGG VFRQVEKVREESRAQSPVEAEFPYGQLS								
TQAPIFISFFIALREMANLPVPSLQTGGL   QDLTVSDPYILPLAVTATMWAVLELG   QDLTVSDPYILPLAVTATMWAVLELG   VQSSLQ WMRNVRMMPLITLPITMIFF   FMYWLSSNLFSLVQVSCLRIPAVRTVIL   VVHDLDKLPPREGFLESFKKGWKNAEN   LREREQRMRNQLELAARGPLRQTFTHN   PGKDNPPNIPSSSSSSKPKSK YPWHDT   THOTRETLSSGSTNRLPLISSSATMPS   HRSPNGLFRQSPVK/TPPIPMSFQPVPG   PRGSGNPPHGTSILTAPPALLPHPPTHFT   LIQENNITNHTHSHTHTYTETLSFFLYIG   DRMEWGKSVF   LKRKLSSLNSEVSTIQNTRMLAFKATA   GCTWCLGLLQVQFAAQVMAYLFTINS   FIFLVYCLLSQQVQKQYQK WFREIVKS   ETYTLSSGMPPSKPDSEGDVPFTISE   S65   1915   A 4288   83   406   RNSRPLWCSPPASQPRQAPVSQSCCEP   PSALLAPTKPRALGTLRLYECSPELGT   PAWLLMLCQAPRPQDPDRLTQPEKSL   GQTGASRTPRT   GVTYLTLTFTCSTPWAESSTKFRHHMY   GLTSFGEKVYEELNRLGMMDLSYASD   VLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVY   DHFDHIRAVIGSEFIGIGGNYDGTGRFPQ   DVSTYPVLIEELLSRSWSEELQGVLRG   VFRQVEKVREESRAQSPVEAEFPYGQLS   FHLGASEWTPRLLIWR   FALGATER								
QDLTVSDPIYILPLAVTATMWAVLELG VQSSDLQWMRNVIRMMPLITIPHTHH PMYWLSSNLPSLVQVSCLRIPAVRTVILI VVHDLDKLPPREGFLESFKKGWKNAEI LREREQRMRNQLELAARGPLRQTFTHN PGKDNPPNIPSSISSSSSKPKSKYPWHDT  563 1913 A 4265 623 116 MGGLAPTQTLEPTREYQNTQLSVSYLI THGTRRTLSSGPSNNLPLPLSSSATMPSI HRSPNGGLFRQSPVK/TPPPMSFQPVPG PRGSGNPPHGTSILTAPPALLPHPPTHFT LIQENNNTNHTHSHTHTYTETLSFFLYIG DRMEWGKSVF  564 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGILQVQPAAQVMAYLFTINS FIFLVYCLLS\QQVQKQYQK WFREIVKS ETYTLSSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPI PPSALLAPTKPRALGTLRLYECSPELGTT PAWLLMLCQAPPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNLLNVPI LLKKNGGIVMVTLSMGVLQCNLLANVI DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIELLSRSWSEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR			,					
VQSSDLQWMRNVIRMMPLITLPITMHF   PMYWLSSNLFSLVQVSCLRIPAVRTVLI   VVHDLDKLPPREGFLESFKKGWKNAE    LREREQRMRNQLELAARGFLRQIFTIN    PGKDNPPNIPSSISSSSSKPKSKYPWHDT    PGKDNPPNIPSSISSSSSKPKSKYPWHDT    PGKDNPPNIPSSISSSSSKPKSKYPWHDT    PGKDNPPNIPSSISSSSSKPKSKYPWHDT    PGKDNPPNIPSSISSSSSKPKSKYPWHDT    THOTRRTLSSGPSNNLPLPLSSSATMPS    HRSPNGGLFRQSPVK/TPPIPMSFQPVPC    PRGSGNPPHGTSILTAPPALLPHPPTHPT    LIQENNNTNHTHSHTHTYTETLSFFLYIC    DRMEWGKSVF    DRMEWGKSVF    S64				-				
FMYWLSSNLFSLVQVSCLRIPAVRITVIL VVHIDLØKLPPREGFLESFKKGWKNAER LREREQRMRNQLELAARGPLRQIFTHN PGKDNPPNIPSSSSSSSSKPKSKYPWHIDT PGKDNPPNIPSSSSSSSKPKSKYPWHIDT HTTHOTRETLSSGPSNNLPLPLSSSATMPSI HRSPNGGLFRQSPVK/TPPIPMSFQVPVG PRGSGNPPHGTSILTAPPALLPHPPTHFI LIQENNNTNHTHSHTHTYTETLSFFLYIG DRMEWGKSVF  564 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGLLQVGPAAQVMAYLFTIINS FIFLVYCLLSQQVQKQYQKWFREIVKS ETYTLSSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPL PPSALLAPTKPRALGTLRLYECSPELGTT PAWLLMLCQAPRPQDPDRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLITLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNILNVPI LLKKNGGIVMVTLSMGVLQCNLLANVSDHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS PHLGASEWTPRLLIWR								
VVHDLDKLPPREGFLESFKKGWKNAEI LREREQRMRNQLELAARGPLRQTFTHN PGKDNPPNIPSSISSSSKPKSKYPWHDT FORDPNIPSSISSSSKPKSKYPWHDT THOTRETLESGPSNNLPLPLSSSATMPSI HRSPNGGLFRQSPVK/TPPIPMSFQPVPG PRGSGNPPHGTSILTAPPALLPHPPTHPT LIQENNNTNHTHSHTHTYTETLSFFLYIG DRMEWGKSVF  S64 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGLLQVGPAAQVMAYLFTIINS FIFLVYCLLS\QQVQKQYQKWFREIVKS ETYILSKMGPDSKPSEGDVFPTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPI PPSALLAPTKPRALGTLRLYECSPELGT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS. GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMDLSYASD VLEVSQAPVIFSHSAARAVCDNLLNVPI LLKKNGGIVMVTLSMGVLQCNLLANVS DHFDHRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR								
LREREQRMRNQLELAARGPLRQTFTHN PGKDNPPNIPSSISSSSSKPKSK YPWHIDT  563 1913 A 4265 623 116 MGGLAPTQTLEPTREYQNTQLSVSYLI THGTRRTLSSGPSNNLPLPLSSSATMPSI HRSPNGGLFRQSPVK/TPPIPMSFQPVPG PRGSGNPPHGTSILTAPPALLPHPPTHPT LIQENNNTNHTHSHTHTYTETLSFFLYIG DRMEWGKSVF LKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGLLQVGPAAQVMAYLFTIINS FIFLVYCLLS\QQVQKQYQKWFREIVKS ETYTLSSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPI PPSALLAPTKPRALGTLRLYECSPELGT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMDLSYASD VLEVSQAPVIFSHSAARAVCDNLLNVPI LLKKNGGIVMVTLSMGVLQCNLLANVS DHFDHRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR					, ,			-
PGKDNPPNIPSSSSSSSKPKSKYPWHDT   FGKDNPPNIPSSSSSSSKPKSKYPWHDT   FGKDNPPNIPSSSSSSSKPKSKYPWHDT   FGKGNPPHIPSSSSSSSKPKSKYPWHDT   FGKGNPPHIPTRETYNTQLSVSYLL   THGTRRTLSSGPSNNLPLPLSSSATMPSI   HRSPNGGLFRQSPVK/TPPPMSFQPVPG   PRGSGNPPHGTSILTAPPALLPHPPTHPT   LIQENNNTNHTHSHTHTYTETLSFFLYIG   DRMEWGKSVF   LIQENNTNHTHSHTHTYTETLSFFLYIG   DRMEWGKSVF   GCTWCLGLLQVGPAAQVMAYLFTIINS   FIFLVYCLLS\QQVQKQYQKWFREIVKS   ETYTLSSKMGPDSKPSEGDVFPRTSE   ETYTLSSKMGPDSKPSEGDVFPRTSE   FFLYPCLLS\QQVQKQYQKWFREIVKS   GTYSLLSVLRSVGPRQAPVSQSCCCPL   PPSALLAPTKPRALGTLRLYECSPELCT   PAWLLMLCQAPRPQDPDPRLTQPEKSL   GQTGASRTPRT   LNSSQKLACLIGVEGGHSLDSSLSVLRSVGPKYLTLTFTCSTPWAESSTKFRHHMY   GLTSFGEKVVEELNRLGMMIDLSYASD   VLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVVFSVEXSFSRQSPVEXEFPYGQLSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLANVSVLEVSGARAVCDNLANVSVLEVSGARAVCDNLANVSVLEVSGARAVCDNLANVSVLEVSGARAVCDNLANVSVLEVSGARAVCDNLANVS	<b>LEMTRQ</b>	VVHDLDKLPPREGFLESFKKGWKNAEN						
PGKDNPPNIPSSSSSSSKPKSKYPWHDT   FGKDNPPNIPSSSSSSSKPKSKYPWHDT   FGKDNPPNIPSSSSSSSKPKSKYPWHDT   FGKGNPPHIPSSSSSSSKPKSKYPWHDT   FGKGNPPHIPTRETYNTQLSVSYLL   THGTRRTLSSGPSNNLPLPLSSSATMPSI   HRSPNGGLFRQSPVK/TPPPMSFQPVPG   PRGSGNPPHGTSILTAPPALLPHPPTHPT   LIQENNNTNHTHSHTHTYTETLSFFLYIG   DRMEWGKSVF   LIQENNTNHTHSHTHTYTETLSFFLYIG   DRMEWGKSVF   GCTWCLGLLQVGPAAQVMAYLFTIINS   FIFLVYCLLS\QQVQKQYQKWFREIVKS   ETYTLSSKMGPDSKPSEGDVFPRTSE   ETYTLSSKMGPDSKPSEGDVFPRTSE   FFLYPCLLS\QQVQKQYQKWFREIVKS   GTYSLLSVLRSVGPRQAPVSQSCCCPL   PPSALLAPTKPRALGTLRLYECSPELCT   PAWLLMLCQAPRPQDPDPRLTQPEKSL   GQTGASRTPRT   LNSSQKLACLIGVEGGHSLDSSLSVLRSVGPKYLTLTFTCSTPWAESSTKFRHHMY   GLTSFGEKVVEELNRLGMMIDLSYASD   VLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LLKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVIFSHSAARAVCDNLLNVPI   LKKNGGIVMVTLSMGVLQCNLLANVSVLEVSQAPVVFSVEXSFSRQSPVEXEFPYGQLSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLLANVSVLEVSGARAVCDNLANVSVLEVSGARAVCDNLANVSVLEVSGARAVCDNLANVSVLEVSGARAVCDNLANVSVLEVSGARAVCDNLANVSVLEVSGARAVCDNLANVS	HNPLLC	LREREQRMRNQLELAARGPLRQTFTHN						
THGTRRTLSSGPSNNLPLPLSSSATMPSI HRSPNGGLFRQSPVK/TPPIPMSFQPVPG PRGSGNPPHGTSILTAPPALLPHPPTHPT LIQENNNTNHTHSHTHTYTETLSFFLYIG DRMEWGKSVF  564 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGILQVGPAAQVMAYLFTIINS PIFLVYCLLS\QQVQKQYQK WFREIVKS ETYTLSSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPL PPSALLAPTKPRALGTLRLYECSPELGTT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS. GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNLLNVPI LLKKNGGIVMVTLSMGVLQCNLLANV. DHFDHIRAVIGSEFIGIGGNYDGTGRFV DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR	DTLG	PGKDNPPNIPSS\SSSSSKPKSKYPWHDT						
HRSPNGGLFRQSPVK/TPPIPMSFQPVPC PRGSGNPPHGTSILTAPPALLPHPPTHPT LIQENNNTNHTHSHTHTYTETLSFFLYIG DRMEWGKSVF  564 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGLLQVGPAAQVMAYLFTIINS FIFLVYCLLS\QQVQKQYQK WFREIVKS ETYTLSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPI PPSALLAPTKPRALGTLRLYECSPELCTT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELINRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNLLNVPI LLKKNGGIVMVTLSMGVLQCNLLANVS DHFDHIRAVIGSEFIGIGGNYDGTGRFPG DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR	LLPEQN	MGGLAPTQTLEPT\REYQNTQLSVSYLL	116	623	4265	A	1913	563
PRGSGNPPHGTSILTAPPALLPHPPTHPT LIQENNNTNHTHSHTHTYTETLSFFLYIG DRMEWGKSVF  564 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGLLQVGPAAQVMAYLFTIINS FIFLVYCLLS\QQVQKQYQK WFREIVKS ETYTLSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPL PPSALLAPTKPRALGTLRLYECSPELGTT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELINRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNLLNVPI LLKKNGGIVMVTLSMGVLQCNLLANVS DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR	PSMQCK	THOTRRTLSSGPSNNLPLPLSSSATMPSN						
PRGSGNPPHGTSILTAPPALLPHPPTHPT LIQENNNTNHTHSHTHTYTETLSFFLYIG DRMEWGKSVF  564 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGLLQVGPAAQVMAYLFTIINS FIFLVYCLLS\QQVQKQYQK WFREIVKS ETYTLSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPL PPSALLAPTKPRALGTLRLYECSPELGTT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELINRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNLLNVPI LLKKNGGIVMVTLSMGVLQCNLLANVS DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR	<b>PGĠV\L</b>	HRSPNGGLFRQSPVK/TPPIPMSFQPVPG			i			
564 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGLLQVGPAAQVMAYLFTIINS FIFLVYCLLSIQQVQKQYQK WFREIVKS ETYTLSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPL PPSALLAPTKPRALGTLRLYECSPELCT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNILINVPI LLKKNGGIVMVTLSMGVLQCNLLANVS DHFDHIRAVIGSEFIGIGGNYDGTGRFFG DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR	<b>IPTQQSF</b>	PRGSGNPPHGTSILTAPPALLPHPPTHPT						
564 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATA GCTWCLGLLQVGPAAQVMAYLFTIINS FIFLVYCLLS\QQVQKQYQKWFREIVKS ETYTLSSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPL PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNILNVPI LLKKNGGIVMVTLSMGVLQCNLLANVI DHFDHIRAVIGSEFIGIGGNYDGTGRFFC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR		LIQENNNTNHTHSHTHTYTETLSFFLYIC						
GCTWCLGLLQVOPAAQVMAYLFTIINS FIFLVYCLLSIQQVQKQYQK WFREIVKS ETYTLSSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPI PPSALLAPTKPRALGTLRLYECSPELCT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNLLNVPI LLKKNGGIVMVTLSMGVLQCNLLANVS DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRS WSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR		DRMEWGKSVF						
FIFLVYCLLS\QQVQKQYQKWFREIVKS ETYTLSSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPL PPSALLAPTKPRALGTLRLYECSPELCT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEFLNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNILNVPI LLKKNGGIVMVTLSMGVLQCNLLANVI DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR	TAQLFII	ILKRKLSSLNSEVSTIQNTRMLAFKATA	368	3	4270	A	1914	564
FIFLVYCLLS\QQVQKQYQKWFREIVKS ETYTLSSKMGPDSKPSEGDVFPRTSE  565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPL PPSALLAPTKPRALGTLRLYECSPELCT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEFLNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNILNVPI LLKKNGGIVMVTLSMGVLQCNLLANVI DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR	NSLQGF	GCTWCLGLLQVGPAAQVMAYLFTIINS						
565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPL PPSALLAPTKPRALGTLRLYECSPELCT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNLLNVPI LLKKNGGIVMVTLSMGVLQCNLLANV DHFDHRAVIGSEFIGIGGNYDGTGRFP DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR	KSKŠES	FIFLVYCLLS\QQVQKQYQKWFREIVKS			1			
PPSALLAPTKPRALGTLRLYECSPELCT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566  1916  A 4298  1041  229  LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNILINVPI LLKKNGGIVMVTLSMGVLQCNLLANVS DHYDHIRAVIGSEFIGIGGNYDGTGRFPG DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR								
PPSALLAPTKPRALGTLRLYECSPELCTT PAWLLMLCQAPRPQDPDPRLTQPEKSL GQTGASRTPRT  566  1916  A 4298  1041  229  LNSSQKLACLIGVEGGHSLDSSLSVLRS GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNLLNVPI LLKKNGGIVMVTLSMGVLQCNLLANVS DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRS WSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR	PLPSSS	RNSRPLWCSPPASQPRQAPVSQSCCCPL	406	83	4288	Α	1915	565
566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS. GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELINELGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNILINVPI LLKKNGGIVMVTLSMGVLQCNLLANV. DHYDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR		PPSALLAPTKPRALGTLRLYECSPELGT						.
566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS. GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELINELGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNILINVPI LLKKNGGIVMVTLSMGVLQCNLLANV. DHYDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR	SLOEAP	PAWLLMLCOAPRPODPDPRLTOPEKSL(						1
566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRS. GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNILINVPI LLKKNGGIVMVTLSMGVLQCNLLANVI DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR		GOTGASRTPRT			1			
GVRYLTLTFTCSTPWAESSTKFRHHMY GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNILLNVPI LLKKNGGIVMVTLSMGVLQCNLLANVI DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR	RSFYVL		229	1041	4298	A	1916	566
GLTSFGEKVVEELNRLGMMIDLSYASD VLEVSQAPVIFSHSAARAVCDNILNVPI LLKKNGGIVMVTLSMGVLQCNLLANVI DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR		GVRYLTLTFTCSTPWAESSTKFRHHMY						
VLEVSQAPVIFSHSAARAVCDNILINVPI LLKKNGGIVMVTLSMGVLQCNLLANVI DHFDHIRAVIGSEFIGIGGNYDGTGRFPO DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR			ł					
LLKKNGGIVMVTLSMGVLQCNLLANV DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR								
DHFDHIRAVIGSEFIGIGGNYDGTGRFPC DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR			i		1			
DVSTYPVLIEELLSRSWSEEELQGVLRG VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR								
VFRQVEKVREESRAQSPVEAEFPYGQLS FHLGASEWTPRLLIWR			,					
FHLGASEWTPRLLIWR	איטארטענ	VFROVEK VREESE A OSDVE A SEED OF VERO	i					
	STOTO CE						Ì	
የሚያቸው የሚያለው ነው	DEADER	GATPLGSVGGRTGKMDAATLTYDTLRF	1106	1	4299	A	1917	567
The state of the s			1100	•	7277	**	4	
l l l l l l l l l l l l l l l l l l l		DFPETSEPVWILGRKYSIFTEKDEILSDV					ļ	. !
		WFTYRKNFPAIGGTGPTSDTGWGCMLR						
MIPAQALVCKHLGKDWRWTQRKRQPD	rusyfs	MIFAQALVCRHLGRDWRWTQRKRQPD		ł				
		VLNAFIDRKDSYYSIHQIAQMGVGEGKS		ļ				ļ
		WYGPNTVAQVLKKLAVFDTWSSLAVH	1				ł	[
	DSDRH	NTVVMEEIRRLCRTSVPCAGATAFPADS		1			.	1
	RLGLTD	CNGFPAGAEVTNRPSPWRPLVLLIPLRL		1				.
		INEAYVETLKHCFMMPQSLGVIGGKPNS		ļ				İ
		FIGYVGEELIYLDPHTTQPAVEPTDGCFI		ļ				
	LSTQAF	<b>FHCQHPPCRMSIAELDPSIAVVRGGHLS</b>		i			ĺ	1
GAECCLGMTRKTFGFLRFFFSMLG	-	GAECCLGMTRKTFGFLRFFFSMLG						
568 1918 A 4300 2012 1843 SRKFLTTPIVLYFLTSFYTKYDQIHFVLN			1843	2012	4300	A	1918	568
LMSVLIPKLPQLHGVRIFGINKY	/LNTVS							
569 1919 A 4302 186 531 WTFCLFL/WWVPFSARWI TOGHVKRA	LNTVS	<u> </u>						

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	bod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	I	<b>]</b> .	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	ł		residue of	sequence	V-Tracing V-Illians + Change in
	1	1	i	peptide	Scharter	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=-possible
]		1	1	sequence		nucleotide insertion
	<del> </del>	$\vdash$	<del>                                     </del>	sequence	<del> </del>	
1	1	1	1	1	1	LLHCARLNGRPVCEDSFSQEVRVNVCVSMHI
1	ł					CVWWGVGCVKCLPPRAHHIWQEKPLGPHRT VTESKLEAEGKTKEKAREKERKKKS
570	1920	A	4308	3	869	
7.0	1 ->20	1 **	7300	3	009	RSGQGKVYGLIGRRRFQQMDVLEGLNLLITIS
ł	1	1	1	ŀ		GKRNKLRVYYLSWLRNKILHNDPEVEKKQG
	1		1			WTTVGDMEGCGHYRVVKYERIKFLVIALKSS
	1	l				VEVYAWAPKPYHKFMAFKSFADLPHRPLLV
	1	1				DLTVEEGQRLKVIYGSSAGFHAVDVDSGNSY
	· ·					DIYIPVHIQSQITPHAIIFLPNTDGMEMLLCYE
	[	1	[			DEGVYVNTYGRIKDVVLQWGEMPTSVAYIC
	1		i .		]	SNQIMGWGEKAIEIRSVETGHLDGVFMHKRA
/		1				QRLKFLCERNDKVFFASVRSGGSSQVYFMTL
571	1921	A	4309	9	524	NRNCIMNW
	.,	1 13	1 4305	3	324	ASREMDVTKVCGEMRYQLNKTNMEKDEAE
1						KEHREFRAKTNRDLEIKDQEIEKLRIELDESK
			1			QHLEQEQQKAALAREECLRLTELLGESEHQL
		ľ				HLTRQEKDSIQQSFSKEAKAQALQAQQREQE
J	J l		1			LTQKIQQMEAQHDKTENEQYLLLTSQNTFLT
572	1922	Α	4318	1	1119	KLKEECCTLAKKLEQISQ
3,2		A.	4516	•	1119	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE
	1		l 1			DFPETSEPVWILGRKYSIFTEKDEILSDVASRL
	1		l I			WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ
Í						MIFAQALVCRHLGRDWRWTQRKRQPDSYFS
	1					VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ
	1					WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD
	1					NTVVMEEIRRLCRTSVPCAGATAFPADSDRH
1	1 1		1			CNGFPAGAEVTNRPSPWRPLVLLIPLRLGL\T DINEAYV\EIL\KHCFHGWPQFPG/VVHREGK
						PNSAHYFIGYVGEELIYLDPHTTQPAVEPTDG
	i I					CFIPDESFHCQHPPCRMSIAELDPSIAVVRGGH
	1 1					LSTQAFGAECCLGMTRKTFGFLRFFFSMLG
573	1923	A	4333	363	1066	GGVPVGLASKPFQILYGHTNEVLSVGISTELD
			,,,,,,	300		MAVSGSRDGTVIIHTTQKGQYMRTLRPPCESS
			ì			LFLTIPNLAISWEGHIVVYSSTEEKTTLKVERM
			İ			HYICFSINGKYLGSQILKEQVSDICIIGEHIVTG
[	l i					SIQGFLSIRDLHSLNLSINPLAMRLPIHCVCVT
1		1	·	ì		KEYSHILVGLEDGKLIVVGVGKPAEVKPSISN
1	[	- 1	.			FISHAVGDYFGSPSFQLIEKSPLGINKLKAKFD
1		- 1				PSKGSK
574	1924	A	4346	359	1234	MDTLEEVTWANGSTALPPPLAPNISVPHRCLL
1		ł				LLYEDIGTSRVRYWDLLLLIPNVLFLIFLLWK
1	} <b> </b>	ļ		I		LPSARAKIRITSSPIFITFYILVFVVALVGIARA
	<b> </b>			. [		VVSMTVSTSNAATVADKILWEITRFFLLAIEL
I		ļ	1	f	ļ	SVILGLAFGHLESKSSIKRVLAITTVLSLAYSV
ļ	]	- 1	1	ŀ	Ì	TQGTLEILYPDAHLSAEDFNIYGHGGRQFWL
	1	- 4	I	ĺ	ı	VSSCFFFLVYSLVVILPKTPLKERISLPSRRSFY
ł		l				VYAGILALLNLLQGLGSVLLCFDIEGLCCVD
						ATTFLYFSFFAPLIYVAFLRGFFGSEPKILF
575	1925	Ā	4360	2038	1512	GCWWRHPWLASQRDCLDCRIQLAEKFVKAV
		1				SKPSRPDMNPIRVKEVYRLEEMEKIFVRLEM
		- 1	- 1	ŀ	]	KIIKGSSGTPKLSYTGRDDRHFVPMGLYIVRT
		ľ	- 1			VNEPWTMGFSKSFKKKFFYNKKTKDSTFDLP
1			- 1		i	ADSIAPFHICYYGRLFWEWGDGIRVHDSQKP
		J	j		j.	QDQDKLSKEDVLSFIQMHRA
576	1926	A	4365	69	500	QVEGRQGREVKRTAWRISPVWRPARCRRRST
		**	-1505	"	200	POPPE PRODUCTION OF THE COLUMN TO THE COLUMN
			ŀ		ŀ	PQP/PE/PGAQQQERHRQGEAPMQALDPRAEP GPQAQSHAACQPEPEPPRVLLDPTAARGGVQ
				j		CDD/CI CDLIDCI ADLIDCITETTI DOCCOTA DO LO
		ł		- 1		GRP/GLSRHPGLAPHPQTHTPWPQSGRLPCAS EPLPLGGIRPTPGLEPKGRDLM
			1	ľ		THE LANGING LANGING LANGEST AND THE PROPERTY OF THE PROPERTY O

CEO III	Lecom	137.4	Tero	1 10 11		
SEQ ID NO: of	SEQID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	nou	in NO.	nucleotide	nucleotide location	D-Aspartic Acid, E-Ghrtamic Acid, F-Phenylalanine, G-Glycine, H-Histidine,
cotide	seq-	1	USSN	location	corresponding	
seq-	uence	1	09/496	correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ľ	'		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		]	,	peptide	Scquato	
	İ	· ·		sequence		nucleotide insertion
577	1927	A	4366	785	502	SAPPKKKNGVLFLSPRLKSSGAIWVHSTPTLW
						ASSNSRASTPKVAGITGARPHARIIFVFLIEMG
			}			FHNVGQAGL/DTLTLVICPPQPPKLLGLQM
578	1928	A	4367	1	221	FFFFLKKSRCVTQAGVQG\PISLHPPPPGFKRF
				1		SRLSLLSSWDYRHP/HAANFCIFSRDG\VSPYW
	Ì		Į			SGWSRTPDLR
579	1929	A	4383	1	224	FETESHSVTQAGMQWHNLGSLQPMP/PGLKR
	<b>{</b>	l	1	ł	ł	FSCLRLQSSWDHRHAPPHLAHFCIFSRDGVSP
		1				CWPGWSSTPDLK
580	1930	Α	4397	410	94	SRLKPYSTNVTAKKLPATNIPNLDCFTAKLYO
		Į.		l		\VFKKGI\IHILHELFONKEEGAFPNS/FYEASFT
		1				LRPKSDRDIAKEESYSTISLLSTDTKILMSKYK
				ľ	-	QLKSSDL .
581	1931	Α	4414	670	3	VLVHRQCGGILRLRRKEAVSVLDSADIEVTDS
						RLPHATIVDHRPQHRWLETCNAPPQLIQGKA
		l				RSAPKPSQASGHFSVELVRGYAGFGLTLGGG
		ľ				RDVAGDTPLAVRGLLKDGP\AQRCGRLEVGD
						LVLHINGESTQGLT\HAQAVERIRAGGPQLHL
				!		VIRRPLETHPGKPRGVGEPRKGVVPSWPDRSP
						DPGGPEVTGSRSSSTSLVQHPPSRTTLKKTRG
						SPE
582	1932	A	4424	194	449	VLYIRKKKRLEKLRHQLMPMYNFDPTEEQDE
						LEQELLEHGRDAASVQAATSVQAMQGKTTL
						PS/QGPLQRPSRLVFT/DVANAIHV
583	1933 .	Α	4435	1	166	APGPPVPPPGSPPEQMPGPCPASMPP/DPPPGS
						PPEQMPGPCPVSAPP/GPPPGSPPEQMPGPCPV
<u> </u>	100.4	<b>!</b>				SAPPALLQDTSV
584	1934	Α	4439	1	628	SATPQQPSAPQHQGTLNQPPVPGMDESMSYQ
		1				APPQQLPSAQPPQPSNPPHGAHTLNSGPQPGT
		l				APATQHSQAGPATGQAYGPHTYTEPAKPKK
		l				GQQLWNRMKPAPGT\EVSSSTSRSDPLLLPPR
		1	1			ALAPTQRASTVVLAPSPT/SEKVQNHSGSSAR
						GNLSGKPDDWP/LGHERVCGALLHRL*VGGG
585	1935	A	4463	10	144	QGPHGKAAQGGAAGAAGRLGLYH
303	1933	Α .	4403	10	144	HKPVTNSRDTQEVPLEKAKQVLKIIATFKHTT
586	1936	A	4464	1309	103	SIFDDFAHYEKRQ
-500	1930	^	4404	1303	103	LNAESYVSFTTKLDIPTAAKYEYGVPLQTSDS
		l				FLRFPSSLTSSLCTDNNPAAFLVNQAVKCTRK
						INLEQCEEIEALSMAFYSSPEILRVPDSRKKVPI TVQSIVIQSLNKTLTRREDTDVLQPTLVNAGH
					. !	FSLCVNVVLEVKYSLTYTDAGEVTKADLSFV
		l	'			LGTVSSVVVPLQQKFEIHFLQENTQPVPLSGN
Į		l	1			PGYVVGLPLAAGFOPHKGSGIIOTTNRYGOLT
						ILHSTTEODCLALEGYRTPVLFGYTMOSGCK
						LRLTGALPCQLVAQKVKSLLWGQGFPDYVA
						PFGNSQGP/ADMLDWVPIHFITOSFNRKDSCO
		I				LPGALVIEVKWTKYGSLLNPQAKIVNVTANLI
ì		ŀ				SSSFPEANSGNERTILISTAVTFVDVSAPAEAG
		l				FRAPPAINARLPFNFFFPFV
587	1937	A	4471	614	387	LLGRASAC/LQLQSSW/D/HRPMLPYLANFVF
1	'	1				CKDR/SFTWLPRLVLNSWLQVILLPWPPTGCD
<u> </u>	1	1	1			NKHEPPCPATKRRHSGSI
588	1938	A	4480	1720	1458	HDLGSLOPPPPGFKRFSCLSLPSSWDYRLMPP
- 1	-	l				CPANFCIIII/DFLVETGFHHVGQASHELLTSGD
		Į.		,		PPTSASQSAGITGMSYHTWFGES
589	1939	Ā	4487	922	332	APVITSPRVGOPW/RTALALRSLYRARPSLRC
202					[	
109		ŀ				PPVELPWAPRRGHRLSPADDELYORTRISTIO
369						PPVELPWAPRRGHRLSPADDELYQRTRISLLQ REAAQAMYIDSYNSRGFMINGNRVLGPCALL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D-Aspartic Acid, E-Glutamic Acid,
nucl- eotide	peptide	1	in USSN	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq- uence		09/496	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	ucilce	1		correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uciice	ļ	İ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l	ŀ	ł	l	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ	ļ		peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		Ì		sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
	<del> </del>		ļ	sequence		
						VVVGTGDRTERLQSQVLQAMRQRGIAVEVQ DTPNACATFNFLCHEGRVTGAALIPPPGGTSL
		İ			<b>!</b>	TSLGQAAQ
590	1940	A	4492	1	472	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT
""	.,,,,	l ^ <b>`</b>	1772	1 *	472	PFSCLSLPSSWDYRRPPLRPANFFVFLVETGFP
ł		ł	ł			RFSRDGLDLLT/S/GDPPTSASQSAGITGVSHR
	İ	•				ARPKRIGEPRRKCGNAVVWPSTSLGDHRVTS
						VPHQGGLPGPIRVAPSSAGQREASQGPPGR
591	1941	Α	4495	1444	1116	IAARFTLAKTWNQLKRP\TMIDSIKKTR\YIYT
						MEYYADTERNEIMSF\AGTWVELEAIILSKLM
						LKDNWVEDTIPQGAVPCTATAEGMKRLLFAL
						EPWDSSCFPHPSSGV
592	1942	Α	4496	2	919	RTRPLFSGRPTRPVCTMSDERRLPGSAVGWL
						VCGGLSLLANAWGILSVGAKOKKWKPLEFL
						LCTLAATHMLNVAVPIATYSVVQLRRQRPDF
						EWNEGLCKVFVSTFYTLTLATCFSVTSLSYHR
						MWMVCWPVNYRLSNAKKQAGHTVMGIWM
						GSFILSALPAVGWHDTSERFYTHGCRFIVAEI
						GLGFGVCFLLLVGGSVAMGVICTAIALFQTL
						AVQVGRQADHRAFTVPTIVVEDAQGKRRSSI
						DGSEPAKTSLQTTGLVTTTVFTYDCLMGFPVL
593	1943		1606		100	GPFSLADTHLSDLPYTWGDRDSGGACVM
393	1943	Α	4506	2	193	FFFEAESCSVPQAGVQRPDLGWLHAPPP\GSC
				}		HFPASASQVAGTTHARHHTQLIF\AFLVENGL
594	1944	A	4507	1327	647	C KMAGGVRPLRGLRALCRVLLFLSQFCILSGG
3,4	1777	Λ	4307	1327	047	ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS
						CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII
				_ [		NMTCRFCWQLPETDYECTNSTSCMTVSCPRO
						RYPANCTVR\DHVHCLGNRTFPKMLYCNWT
		ļ				GGYKWVYGLWLLRHHPRWGLGADRF\YLGP
İ	i	I				VAGTASGKLFSFGGLGIWILIDVLLIGVGYVG
						PADGSLYI
595	1945	Α	4512	533	264	FFFKMESYSVARLECSGAISAPCNLHLLGSNN
				l		SPASASRV/AGNIGARHHTQQIFVLLVQMRVH
						YVGQDGLDLL/NLMIHPPRSPKVLGLQA
596	1946	A	4513	3	1674	HASDHLYPNFLVNELILKQKQRFEEKRFKLD
	•		l	1		HSVSSTNGHRWQIFQDWLGTDQDNLDLANV
	ſ	l	1	İ		NLMLELLVQKKKQLEAESHAAOLOILMEFLK
						VARRNKREQLEQIQKELSVLEEDIKRVEEMS
		ļ			- 0	GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSOP
						PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE
		- 1				DLEQCYFSTRMSRISDDSRTASQLDEFQECILS
		J	1		2	KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV
}	ł	ł	1		ļ	FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA
İ	į	i				VDIHYPENEMTCNSKISCISWSSYHKNLLASS
		1			1	DYEGTVILWDGFTGQRSKVYQEHEKRCWSV
	Į	l			1	DFNLMDPKLLASGSDDAKVKLWSTNLDNSV
	l	ı		ļ		ASIEAKANVCCVKFSPSSRYHLAFGCADHCV
	l	l	j	[		HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG
		l	- 1			EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN
	İ		l	İ		EKNFV\GLASNGDYIACGSENNSLYLYYKGLS
I	l	l	ł	1	1	KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV
507	1047		4510	-535	204	CWRALPDGESNVLIAANS\QGTT\KVLELV
597	1947	A	4518	536	824	RSLALSPGLECSGMISAHCNLHLLGSSDPPTS
ł				İ		ASQVAEITSVRHHTWLIFCILGQMGFHHVGE
598	1040			<b></b>		QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP
J70						
1	1948	A	4524	1	384	FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF TLGKLPRKTLSVKLMKNRDEVQAMIYDDGSS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  RRREMQSQSVMLALRRGDAVWLLSHDHDG YGAYSNHGKYITFSGFLVYPDLAPAAPPGLG ASELL
599	1949	A	4526	366	776	MGQPAPYAEGPIQGGDAGELCKCDFLVFTSP NPEAVCEAGTPAMFQTAWRQMESCSI/AQAG VQWRDPGSLHPPPLGFKRFSCLSLPSSWDYK HAPPHPANFCIFSRDQVSPCWPGWSRSLDLVI PPPWLPKVLGLQA
600	1950	A	4529	776	334	FFFETESCYVAQAGVQWCDLCSLQAPPPG\SS DPPASASRVAGTTGARHHTQLIFVFLVETGFH \MLARDGLKLLTSSDPPASASQSSWDYRREPP RLANFFVFLVETGSRYVAQAGVQWLFTGAIP LLISTGVLTCSVSDLGRFTPP
601	1951	A	4533	1460	403	HEVQESIHFLESEFSRGISDNYTLALITYALSS VGSPKAKEALNMLTWRAEQEGGMQFWVSSE SKLSDSWQPRSLDIEVAAYALLSHFLQFQTSE GIPIMRWLSRQRNSLGGFASTQDTTVALKALS EFAALMNTERTNIQVIVTGPSSPSPVKFLIDT HNRLLLQTAELADGTANGSV/SISANGFGFAI CQLNVYYNVKASGSSRRRSIQNQEAFDLDV AVKENKDDLNHVDLNVCTSFSGPGRSGMAL MEVNLLSGFMVPSEAISLSETVKKVEYDHGK LNLYLDSVNETQFCVNIPAVRNFKVSNTQDA SVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSD VQRLPSL
602	1952	A	4540	1963	295	MRAPGRPALRPLPLPPLLLLLLSSPWGRAVPC VSGGLPKPANITFLSINMKNVLQWTPPEGLQG VKVTYTVQYFIYGQKKWLNKSECRNINRTYC DLSAETSDYBHQYYAKVKAIWGTKCSKWAE SGRFYPFLETQIGPPEVALTTDEKSISVVLTAP EKWKRNPEDLPVSMQQIYSNLKYNVSVLNT KSNRTWSQCVTNHTLVLTWLEPNTLYCVHV ESFVPGPPRRAQPSEKQCARTLKDQSSEFKAK IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HPANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCB PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN
603	1953	A	4543	3	600	YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN YKVIGHEDLETAV
604	1954	A	4548	3	938	QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW STAGDPPIPYLTTYGQLSNGDHHFMHDAVFO QPGGLGNNIYQHRFNFFPENPAFSAWGTSGS QGQQTQSSAYGSSYTYPPSSLGGTVVDGQPG FHSDTLSKAPGMNSLEQGMVGLKIGDVSSSA VKTVGSVVSSVALTGVLSGNGGTNVNMPVS KPTSWAAIASKPAKPQPKMKTKSGPVMGGG LPPPPIKHNMDIGTWDNKGPVPKAPVPQQAP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	)	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	1	ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	I	1		residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ					/=possible nucleotide deletion, \=possible nucleotide insertion
	<del> </del>	<del></del> -		sequence	<del> </del>	
						SPQAAPQPQQVAQPLPAQPPALAQPQYQSPQ QPPQ
605	1955	Α	4553	2	2304	ILLQEKRNCLLMQLEEATRLTSYLQSQLKSLC
						ASTLTVSSGSSRGSLASSRGSLASSRGSLSSVS
		· .				FTDIYGLPQYEKPDAEGSQLLRFDLIPFDSLGR
		ſ				DAPPSEPPGPSGFHKQRRSLDTPQSLASLSSRS
		ĺ				SLSSLSPPSSPLDTPFLPASRDSPLAQLADSCE
		}				GPGLGALDRLRAHASAMGDEDLPGMAALQP
						HGVPGDGEGPHERGPPPASAPVGGTVTLRED
						SAKRLERRARRISACLSDYSLASDSGVFEPLT
[.	·		. 1	İ	İ	KRNEDAEEPAYGDTASNGDPQIHVGLLRDSG SECLLVHVLQLKNPAGLAVKEDCKVHIRVYL
				_		PPLDSGTPNTYCSKALEFQVPLVFNEVFRIPV
				- 50		HSSALTLKSLQLYVCSVTPQLQEELLGIAQIN
						LADYDSLSEMQLRWHSVQVFTS\LNHQGRGR
						LGVQERAPPGTLHTPSPSPA/STDAVTVLLAR
1 1			\			TTAQLQAVERELAEERAKLEYTEEEVLEMER
1						KEEQAEAISERSWQADSVDSGCSNCTQTSPPY
						PEPCCMGIDSILGHPFAAQAGPYSPEKFQPSPL
1						KVDKETNTEDLFLEEAASLVKERPSRRARGSP
					·	FVRSGTIVRSQTFSPGARSQYVCRLYRSDSDS
						STLPRKSPFVRNTLERRTLRYKQSCRSSLAEL MARTSLDLELDLQASRTRQRQLNEELCALRE
ļ ļ		· l	}			LRQRLEDAQLRGQTDLPPWVLRDERLRGLLR
						EAERQTRQTKLDYRHEQAAEKMLKKASKEI
]			ļ			YQLRGQSHKEPIQVQTFREKIAFFTRPRINIPPL
				·		PADDV
606	1956	A	4555	3429	776	PGSGPGPAPFLAPVAAPVGGISFHLQIGLSREP
	,	]	j	j		VLLLQDSSGDYSLAHVREMACSIVDQKFPEC
			l			GFYGMYDKILLFRHDPTSENILQLVKAASDIQ
		- 1				EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLV\RQGLKCEGCGLNYH
i !		ı			ł	KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT
	}	.				SSAELSTSAPDEPLLQKSPSESFIGREKRSNSO
		- 1	l		J	SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV
]					-	CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA
	1	- 1	j	1		PKVPNNCLGEVTINGDLLSPGAESDVVMEEG
]		I	1		- 20	SDDNDSERNSGLMDDMEEAMVQDAEMAMA
1 1	J	l	]	j		ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP
	ŀ	l	ŀ	1		LMRVVQSVKHTKRKSSTVMKEGWMVHYTS
	Ī		- 1		ì	KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE PLSEILSLEPVKTSALIPNGANPHCFEITTANV
1 1	1	ł	1	. 1	- 1	VYYVGENVVNPSSPSPNNSVLTSGVGADVAR
j l	j	1		·	Ì	MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV
]	· ]	j	. ]	J	•	SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI
[	J	ļ	[	. [	[	VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR
	1	į	ŀ	1		NEVAILQNLHHPGVVNLECMFETPERVFVVM
		ł	}	ļ	ł	EKLHGDMLEMILSSEKGRLPEHITKFLITQILV
		l				ALRHLHFKNIVHCDLKPENVLLASADPFPQV
	Į	l	-			KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL
[	İ	1	i	Ì		RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED
	l	j				EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN
	1	l	1	·	Į	LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL
•	}	ļ	l	j	I	RELECKIGERYITHESDDLRWEKYAGEQGLQ
607	1957	A	4563	1 +	4499	YPTHLINPSASHSDTPETEETEMKALGERVSIL
	.,,,	^		·	T127	SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT
]	J	ļ	J	J	,	TVRCMHLLLEAVPAVAPQTSILDLRFNRIREI
	İ		1	Į	ŀ	QPGAFRILRNLNTLLLNNNQIKRIPSGAFEDL
	ł	1		ŀ	l	ENLKYLYLYKNEIQSIDROAFKGLASLEOLYL

S	EQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	iO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
n	ucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
0	otide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
S	eq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
լա	ence	Į .	ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine
1		ļ	l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1			ĺ	ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
- [		Ì		1	peptide	'	/=possible nucleotide deletion. \=possible
L			. ·	j	sequence	J	nucleotide insertion
							HFNQIETLDPDSFQHLPKLERLFLHNNRITHL
1			ĺ				VPGTFNHLESMKRLRLDSNTLHCDCEILWLA
				1			DLLKTYAESGNAQAAAICEYPRRIQGRSVATI
ı			ĺ	ĺ		1	TPEELNCERPRITSEPQDADVTSGNTVYFTCR
		_	i				AEGNPKPEIIWLRNNNELSMKTDSRLNLLDD
1			l		'		GTLMIQNTQETDQGIYQCMAKNVAGEVKTQ
1			ľ				EVTLRYFGSPARPTFVIQPQNTEVLVGESVTL
1			1	.			ECSATGHPPPRISWTRGDRTPLPVDPRVNITPS
							GGLYIQNVVQGDSGEYACSATNNIDSVHATA
1							FIIVQALPQFTVTPQDRVVIEGQTVDFQCEAK
1	J						GNPPPVIAWTKGGSQLSVDRRHLVLSSGTLRI
1							SGVALHDQGQYECQAVNIIGSQKVVAHLTVO
							PRVTPVFASIPSDTTVEVGANVQLPCSSQGEP
	j						EPAITWNKDGVQVTESGKFHISPEGFLTINDV
	İ						GPADAGRYECVARNTIGSASVSMVLSVNVPD
1							VSRNGDPFVATSIVEALATVDRAINSTRTHLF
1		j			ſ		DSRPRSPNDLLALFRYPRDPYTVEQARAGEIF
1	- 1						ERTLQLIQEHVQHGLMVDLNGTSYHYNDLVS
1	ŀ	ł					PQYLNLIANLSGCTAHRRVNNCSDMCFHQKY
	ļ			}			RTHDGTCNNLQHPMWGASLTAFERLLKSVY
	1						ENGFNTPRGINPHRLYNGHALPMPRLVSTTLI
1	-	1		ļ			GTETVTPDEQFTHMLMQWGQFLDHDLDSTV
Ì				ļ			VALSQARFSDGQHCSNVCSNDPPCFSVMIPPN
1	- 1			f			DSRARSGARCMFFVRSSPVCGSGMTSLLMNS
1		·		į	1		VYPREQINQLTSYIDASNVYGSTEHEARSIRD
	- 1		İ		1		LASHRGLLRQGIVQRSGKPLLPFATGPPTECM
	- 1	1					RDENESPIPCFLAGDHRANEQLGLTSMHTLW
1	1	ļ	ļ	1	J		FREHNRIATELLKLNPHWDGDTTYYETRKIVG
1	- 2	- 1			ł	ļ	AEIQHITYQHWLPKILGEVGMRTLGEYHGYD
1	- 1	l	- 1	1	1		PGINAGIFNAFAT\AAFRFGHTLVNPLLLPGLD
		[			ŀ		ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLL
1	ŀ			1			RGLFGVAGKMRVPSQLLNTELTERLFSMAHT VALDLAAINIQRGRDHGIPPYHDYRVYCNLS
1	. 1	ł	- 1	-	}		AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID
	i						LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLR
1	i	1	i		ļ		DGDRLWYENPGVFSPAQLTQIKQTSLARILCD
l	- 1	}	- 1	1	Ī	ł	NADNITRVQSDVFRVAEFPHGYGSCDEIPRVD
1	- 1		İ			1	LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLE
1	- 1						PSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS
1	- 1	- 1	- 1	Į.	.)	ļ	A\FSTRSDASG\TNDFQRVCSWEMQKTITDLR
ŀ	ı	1	1		1		TQIKKLESRUSTTECVDAGGESHANNTKWK
[	- 1	i	1		ł		KDACTICECKDGQVTCFVEACPPATCAVPVNI
L					. 1.	•	PGACCPVCLQKRAEEKP
60	8	1958	A	4566	354	1135	FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAL
1	- 1		1		i		IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKL
i	- 1	1	- 1	1	• 1	l	GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF
I	101		j				YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ
1	- 1	ł	- 1		ŀ	1	LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD
l			- }			l	EFTIAATDDNHIFAWGNGGNGRLAMIPTERP
		- 1	į	}	-	i	HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILI
1	- 1	1	1	- 1	į.	}	VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGE
<u></u>						I	GGPDAW
609	9	1959	A ·	4567	1	412	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT
	- 1	-	- 1	i			PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF
			- 1	- 1			HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH
1	- 1	İ	- 1	1	1	. [	RARPRINLRNVIYSFAVTYCLNYISLAMSSTL
				- 1			KLSFHVLSGS
610		1960	A	4570	697	467	ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT
610	0	1960	A	4570	697		
610	0	1960	A	4570	697		ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT GTCDYAQLIFAFLVEMGFHHVGQDGLHLL\N LVIRPPRPPKVLGLQA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	j	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop coden,
	1			peptide	sequence	/=possible nucleotide deletion. \=possible
	1	1	1	sequence	1	nucleotide insertion
611	1961	A	4571	25	1396	
011	1 1201	^	43/1	2	1390	ADPHITVIRFFPAASATKRVLPPVLRVSSPRT
Ì						WNPNVPESPRIPAPRLPKRMSGAPTAGAALM
1	ľ	l	ł		ł	LCAATAVLLSAQGGPVQSKSPRFASWDEMN
· .			1			VLAHGLLQLGQG\CANT\GAHPQSAERAGA\R
	1	l		İ		LSACGSACQGTEGSTDLPLAPESRVDPEVLHS
ļ	ļ	1	J		1	LQTQLKAQNSRIQQLFHKVAQQQRHLEKQHL
		l		1		RIQHLQSQFGLLDHKHLDHEVAKPARRKRLP
						EMAQPVDPAHNVSRLHRLPRDCQELFQVGER
						QSGLFEIQPQGSPPFLVNCKMTSDGGWTVIQR
		1	1			RHDGSVDFNRPWEAYKAGFGDPHGEFWLGL
	1 .	l	}	}		EKVHSITGDRNSRLAVQLRDWDGNAELLQFS
						VHLGGEDTAYSLQLTAPVAGQLGATTVPPSG
	İ			ł	•	LSVPFSTWDQDHDLRRDKNCAKSLSGGWWF
	ļ ·		l	}		GTCSHSNLNGQYFRSIPQQRQKLKKGIFWKT
	1	ł				WRGRYYPLQATTMLIQPMAAEAAS
612	1962	A	4575	162	3	FFFETESRSVAQAGVQWRDLSSLQPPPPG\SR
						GSPASASPVAGITGTRHHRTRG
613	1963	Α	4584	687	321	PLAQRRPFLWVTVKTNGHIWGSSTYPHFWGS
						SNS/PASASQVAGIPNARHQARIIFVFLVEPRF
		l .	1	i		HHVGRAGLGFL/NLAICLPOHPKVLGLOACN
			i	'		LNIKPHPAHKYISMIQFNVHFMCMSVHIYI
614	1964	Α	4589	727	299	PGSAQSAQRGRGRRRARAGSATQITMYSFMG
-		]	1	•=•		GGLFCAWVGTILLVVAMATDHWMQYRLSGS
		ţ		1		FAHQGLWRYCLGNKCYLQTDSIAYWNATRA
						FMILSALCAISGIIMGIMAF/GWVAVLMTFFA
						GIFYMCAYRVHECRRLSTPR
615	1965	Α	4590	2	414	
.013	1505	1 ^	4370	<i>-</i>	414	TILPEKIQAWAQKQCPQSGEEAVALVVHLEK
		1				ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ
						PEQVETQPRAVSREEPGSLHSGHQEQLNRKR
	1					ERRPLPKNARPSPWVPALADEWNTLHQEVTT
616	1966	A	4592	773	488	TRLPAGSQEPVKD
010	1,700	^	4332	113	400	DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL
						SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ
617	1967		4505	04	400	AGLELPTSGDPPALASQSAGITGVTTVPSGPG
017	190/	В	4595	84	478	XRHGLREPLLERRCAAASSFQHSSSLGRELPY
		İ				DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT
						REKQLQELQQQEEEERQRQQRREERRQQNL
					•	RARSREHPVVGHPDPALPPSGVNCSGCGAEL
(10	1000					HCQDAR*
618	1968	A -	4596	2945	1188	ARSRNSARGVYGMCVDTLFLCFLEDLERNDG
	'	ł				SAERPYFMCSTLKKPLARRCFPAIHAYKGVL
		ļ				MVGNETTYEDGHGSRKNITDLVEGAKKANG
			[	- 81	İ	VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA
						MSLLSIILLHLLAGIMGWVMIIMEI\SELGYRIF
	·					HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV
					1	YLHLROTWLAFMIILSILEVIIILLLIFLRKRILI
			1			AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI
			1 1			AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT
						CNPETFPSSNESRQCPNARCQFAFYGGESGYH
		1				RALLGLQIFNAFMFFWLANFVLALGQVTLAG
		,			j	AFASYYWALRKPDDLPAFPLFSAFGRALRYH
					İ	TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN
						KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM
						IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV
			]	l		TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT
						APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC
						VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL
	1				1	
619	1969	A	4601	2	357	LNKTNKKAAES RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	L Amino soid source (A. 41 ' C. C
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			<b> </b>	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1	ĺ	İ	residue of	sequence	V=Turcoine, V=Valine, w=1typtopnan,
1	Ĭ	ł		peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	Ì	l	ł	sequence		/=possible nucleotide deletion,  =possible nucleotide insertion
	<del></del>			Sequence	<del></del>	nucleotide insertion
Į.	i	]	,			GLAYGHFSYEFSNHRDVVVDLQGWVTGNGK
1.	1	l				GLIYLTDPQIHSVDQKVFTTNFGKRGIFYFFN
620	1970	A	4606	1	0415	NQHVECNEICHRLSLTRPSMEKPCKS
020	1970	^	4000	1	2415	MERLWGLFQRAQQLSPRSSQTVYQRVEGPR
	}					KGHLEEEEEDGEEGAETLAHFCPMELRGPEP
i I	ì		1		1	LGSRPRQPNLIPWAAAGRRAAPYLVLTALLIF
1	F	•				TGAFLLGYVAFRGSCQACGDSVLVVSEDVN
1						YEPDLDFHQGRLYWSDLQAMFLQFLGEGRL
						EDTIRQTSLRERVAGSAGMAALTQDIRAALS
1					J	RQKLDHVWTDTHYVGLQFPDPAHPNTLHWV
1						DEAGKVGEQLPLEDPDVYCPYSAIGNVTGEL
						VYAHYGRPEDLQDLRARGVDPVGRLLLVRV
					•	GVISFAQKVTNAQDFGAQGVLIYPEPADFSO
						DPPKPSLSSQQAVYGHVHLGTGDPYTPGFPSF
1 1						NQTQFPPVASSGLPSIPAQPISADIASRLLRKL
						KGPVAPQEWQGSLLGSPYHLGPGPRLRLVVN
						NHRTSTPINNIFGCIEGRSEPDHYVVIGAORDA
j l						WGPGAAKSAVGTAILLELVRTFSSMVSNGFR
1 1						PRRSLLFISWDGGDFGSVGSTEWLEGYLSVL
	1					HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSL
]	]		,		;	IESVLKQVDSPNHSGQTLYEQVVFTN\PSWD\
	ſ					AEVIRPLPM\DSSAY\SFTAFVGVPAVEFSFME\
1 1	1	1	1			DDQ\AYPFLHTKEDTYENLHKVLQGRLPAVA
1 1	ľ	1	i			QAVAQLAGQLLIRLSHDRLLPLDFGRYGDVV
1 1			ł			LRHIGNLNEFSGDLKARGLTLQWVYSARGDY
1 1	- 1	1	}			IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV
1	I	- 1				EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL
1 1	)					DHLRLLRSNSSGTPGATSSTGFQ\ESRFRRQL\
						ALL\TWDACKGAANALSGDVWNIDNNF
621	1971	Α	4610	793	334	ISRVDDFVGSGIANVIIAVAIFSIPAFARLVRG\
1 1	1			1		NILVLKQQTFIESARSIGASDMTVLLRHILPGT
i I		- 1				GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP
1 1		1	1	ŀ		EWGAMLNEARADMVIAPHVAVFPALAIFLTV
1 1	j	. 1	ļ.			LAFNLLGDGLRDALDPKIKG
622	1972	Α.	4614	2	820	LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ
i 1		1		-		CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF
1 1	- 1	J	J	J	ļ	RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK
f l	1					SCVILLGLLLLYDVFFVFITPFITKNGESIMVEL
j l	1	ļ	1		ļ	AAGPEGNNEKNDGNI VE ATGODE ADITOUT TO
	1	1	İ	1	i	AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL
		1	j	İ	1	I IAACDBEDAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
	ľ	i	- 1	. 1	· · · · · · · · · · · · · · · · · · ·	LIAYCRRFDVQTGSSYIYYVSV\TVAYAIGMIL
		i			1	TFVVLGLMKKGQPALLYLVPCTLITA/CQFV
623	1973	A	4619	17	691	AWETVREMKKFWERVTS
	-7.5	••	7017	*′	U21 .	TLVSVVEFVRRADLTREDLAPSSVDSGQAGF
		- 1	1	i	j	GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT
1	1	- 1		f	ŀ	DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ
	1	- 1	- 1	ł	ł	IGHFLCLVILVYCAEYINEAAAMNWRLFSKY
	1	- 1	1	ŀ		QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW
	ł	ı	1	J		KTLNVMTDLKNAQERRKEKKRRRKED*GAA
		- 1	1		Į.	AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG
-	1001		1425			LKNRCFI
624	1974	A	4622	164	668	VSCYTALQSIMNQPESANDPEPLCAVCGQAH
	ļ	1	1	Ì		SLEENHFYSYPEEVDDDLICHICLQALLDPLD
		.1	1		ľ	TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL
		1	į			QHCKKSSILVNKLLNKLLVTCPFREHCTQVL
	- 1	- 1	ł		1	QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED
						CLSPGVHHCSEV
625	1975	A	4625	474	473	CFLSPSPLLPPLLLSSSSSPSFPLPPPPTLLPSTLP
						PPLLIPSS*LSP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible mucleotide deletion, \=possible
626	1976	A	4629	sequence 249	3	nucleotide insertion  KLKGNECFCYHCNVCIFLMIKK*GLFLC*IYFI LFFET*SHSFTRLECSGTISAHCSLQLQGSSNSP ASASQVAGIAGTHH
627	1977	Α	4635	1	301	FFFFETKPFFAPQAGGQGPSRGSLNPLPTGLK QFSGLTLSRSGNNGPRPPPRVNFGILRGNGVP PGGAG*PRPPDLRGPPGLAPPQGGNNGGDPP ARAYL
628	1978	A	4648	1357	782	KLFSSQRLFGPHIQAINPSFLLLSFFPS*LLAMR TVGNNAFILVFLVYRIVLLLF*HV*PAYFQPSK NKTAKINCN*RPFLFLVCYLL*AELHIGIFIANF YDCIPNKLNEHLWPKLLQSLIFHVDFCGFLHK VFYICFTEFLLFLYFL*LFIIKVSCSII*CSTICVF SYKSFAVIIFFVDNTRFFSFGF
629	1979	A	4660	18	999	HHELHTLELLQNPKEVLTRSEIQDVNYSLEAV KVKTVCQPLMKEMLKRFQVAVNLAEDTAH PKLVFSQEGRYVKNTASASSWPVFSSAWNYF AGWRNPQKTAFVERFQHLSCVLGKNVFTSG KHYWEVESRDSLEVAVGVCREDVMGITDRS KMSPDVGIWAIYWSAAGYWPLIGFPGTPTQQ EPALHRVGVYLDRGTGNVSFYSAVDGVHLH TFSCSSVSRLRPFFWLSPLASLVIPPVTDRK*G FSSPDQNSFPVVQLRDTHPWALFCPSCLYPG WSIFWVSLTVPFGICPLCASQEAVPWEVGLA NGDGTGNFPRRFWEIFL
630	1980	A	4669	2	358	FFFFFETESHSVAQAGMQWRNLGSLPAPPPGF TPFFCLSLLNGWDYRRPPPHLANFFVLLVETG FHDVGQDGLDLLTS*STPSASQSAEITGVSHC TRLKKIRFAKGHVEFFFESHVE
631	1981	A	4674	953	614	TPIRGTDDEHEECTVQEYSAGKNTCLRPGAV AHTCNPCTLGGRGRWIT*GSGVQDQPGPTWQ NPVFLERRPRALHSSPGLTTQRILWAQGLWV GAGSTGCSRGPRGEGVFREG
632	1982	A	4678	34	314	RSTHASGMISPSFGFMGHLLRLEFEILPSTPNP *LPSYQGEAAGSSLISHLQTFSPDLKGVYCTFP ASGLAPVPTHWTVSELSRSPVATATFC
633	1983	A	4696	1		RTLGMEGERRASQAPSSGLPAGGANGESPGG GAPFPGSSGSSALLQAEVLDLDEDEDDLEVFS KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP DLKDLFITVDEPESHVTTIETFTTYRITTKTSRG EFDSSEFEVRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERFNDDFIETRRKALHKFL NRIADHPILTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP
634	1984	A	4708	421	158	SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRRKTLQFPCYH
635	1985	A	4709	42	341	YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED
636	1986	A	4721	2	351	EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS RSTTVTVA*LMQKLNLSMNDAYYIVIMKMSS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l	ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
		l ·	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		i		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ		1	peptide		/=possible nucleotide deletion, \=possible
				sequence	1	nucleotide insertion
						ISPNFNSMDQPLDFQRTLGLRSPCYNRVPAQK
			<u> </u>			MYFTTPSNHNAYQVDSVQST
637	1987	A	4726	664	253	NTGLTCSIQRKCGETQLYRREENRLILLLQDH
	i .		1	i	l	LKSESFQVLTLSPRLEFSGLISAHCNLRLPGSS
	ł		1			DSSASSSRAAGITGVHHHAWLIFFFLVETGFL
					ļ	HAG*AGLELLTSGDPPASASRSAGITGVSHHA
						RPRETRFL
638	1988	A	4734	24	592	GGMDSRVSGTTSNGETKPVYPVMEKKEEDG
			!		j	TLERGHWNNKMEFVLSVAGEIIGLGNVWRFP
			1		i	YLCYKNGGGAFFIPYLVFLFTCGIPVFLLETAL
			,			GQYTSQGGVTAWRKICPIFEGIGYASQMIVIL
			] ' !	ł	•	LNVYYIIVLAWALFYLFSSFTIDLPWGGCYHE
-620	1000					WNTEHCMEFQKTNGSLNGTSENATSPVIEFW
639	1989	A	4743	1040	699	QGLTLLPRMECSATITAHCSLELPGSIDLPTSA
				l		S*VARTTGTHHHPWLILVLLL*TWGSYYVAQ
			1 1		ł	AGLELLGSSNLPAAMVSQSAQIIGHDHCAWA
640	1990		1000			TSNHVLYTQEGLRRGKEG
040	1990	A	4771	527	2	GRIDCPHPATVLAQPIFIDACSVLGAYQGAQN
			1			WIRRRPCLPSGCLKMNREIGPLQHSLCCPGWS
,				l .	J	QTPGLKAILLRQPPK*LGLQMESHSCPPAWSA
			i !			MARSRLTATSASQVQAILLPQPPGTTDSCSPS
						PDHEQQPLSWVLPPPQKDMNPREQQVALGP
641	1991	A	4780	16	100	QAAALPWAVWRNDCFPR
V+1	1331	A	4/80	10	473	RPSSQCGGIPTGWKKGLAPELSSELSSPPLPAR
- 1						LQLAASPYFSPSWAECPQPVPAGTHATWCLA
						RVWARMTPPGPAGIPSHPLPPPPPERSVPIPSP
×						FPARDSGSRQGHSTDRYKHTDAPRDAHRRVP
642	1992	A	4798	1	487	QRDTDTGVHTGSGTHTHAHTPPEK GYSFRCDIVDYSRSPTALRMARTCWLYYFSK
			.,,,	•	107	FIELLDTIFFVLRKKNSQVTFLHVFHHTIMPW
						TWWFGVKFAAGGLGTFHALLNTAVHVVMY
						SYYGLSALGPAYQKYLWWKKYLTSLQLVQF
j						VIVAIHISQFFFMEDCKYQFPVFACIIMSYSFM
		. }	[		- 6	FLLLFLH
643	1993	A	4799	2	391	LMAFIEMHISGSLVYLKIKTKIYSYFSMLNFLL
- 1						QEIPLSEILRISSPRDFTNISQGSNPHCFEIITDT
						MVYFVGENNGDSSHNPVLAATGVGLDVAOS
1	ł	1			-	WEKAIRQALMPVTPQASVCTSPGOGKDHSK
			i			Q*ASVCTSPGQGKDHSKQ
644	1994	A	4800	488	101	AYPLFAVHPVHTECVAGVVGRAYLLCALFFL
				l		LSFLGYCKAFRESNKEGAHSSTFWVLLSIFLG
j		J				AVAMLCKEQGITVLVRAATWLGPAFSVCPFP
						SYKDIWGWPCLCGVLHAYIPLLV
645	1995	A	4805	458	126	LLWITVLCQTPARPQSTMIHLGHILFLLLLPV
1					]	AAAQTTPGERSSLPAFYPGTSGSCSGCGSLSL
						PLLAGLVAADAVASLLIVGAVFLCARPRRSP
			l			AQEDGKVYINMPGRG
646	1996	A	4817	47	1033	LQGDTWHLSFLSHFSRLHGGVPGRGLLEGNL
				·		LQPQAPGHDMTSIPFPGDRLLQVDGVILCGLT
l		- 1	- 1	}		HKQAVQCLKGPGQVARLVLERRVPRSTQQC
1	j				. ]	PSANDSMGDERTAVSLVTALPGRPSSCVSVT
			1	1		DGPKF*SSN*KRIANGLGFSFVQMEKESCSHL
1	i	ļ				KSDLVRIKRLFPGHPAEENGAIAAGDIILGRE
	l	1	, 1	1		WEGPRKASSSRCRGSWAMQLSVQAGPSFAS
[	i		- 1	İ	ľ	YYPAAVEVLHLLRGAPQEVTLLLCRPPPGAL
		į				PELEQEWQTPELSADKEFTRATCTDSCTSPIL
		ľ				GSRGQLGGTVPPQMQGKAWGLRPESSQKAIR
647	1997	A	4854	1044	335	EGTMGAKTERDLGPVP PRVRGDWPLEKKKSNSNIHPIPSWCGSTDSKD

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	""	in NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		]	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		[·		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/-possible nucleotide deletion, \-possible
		ļ		sequence		nucleotide insertion
	ŀ	ĺ	1			IVMPTYDLTDSVLETMGRVSLDMMSVQANT
	İ	1	[	ſ	ĺ	GPPWESKNSTAVWRGRDSRKERLELVKLSRK
		}				HPELIDAAFTNFFFFKHDENLYGPIVKHISFFD
						FFKHKYQINIDGTVAAYRLPYLLVGDSVVLK
	]					QDSIYYEHFYNELQPWKHYIPVKSNLSDLLEK LKWAKDHDEEAKKIAKAGQEFARNNLMGD
			·			DIFCYYFQTFPRNMPIYK
648	1998	Α	4867	2030	837	AGMLPAVGSADEEEDPAEEDCPELVPMETTO
						SEEEKSGLGAKIPVTIITGYLGAGKTTLLNYI
						LTEQHSKRVAVILNEFGEGSALEKSLAVSQG
						GELYEEWLELRNGCLCCSVKDNGLRAIENLM
						QKKGKFDYILLETTGLADPGAVASMFWVDA
1					ļ	ELGSDIYLDGIITIVDSKYGLKHLAEEKPDGLI
						NEATRQVALADAILINKTDLVPEEDVKKLRT
						TIRSINGLEQUILETQRSRVDLSNVLDLHAFDSL
						SGISLQKKLQHVPGTQPHLDQSIVTTTFDVPG NAKEEHLNMFIQNLLWEKNVRNKDNHCMEV
				•		IRLKGLVSIKDKSQQVIVQGVHELYDLEETPV
	•			:		SWKDDTERTNRLVLLGRNLDKDILKQLFIAT
						VTETEKOWTTHFKEDOVCT
649	1999	A	4873	226	189	DGVSLLLPKLGVQWAQYWAHWQPPLPGFKR
l ·						FSCLSLRSSWD*KCAPPHPAFVFLVEMGFHRV
						GQAGLELRTSGDPPASASQSAGITGVSHLA*P
650	2000		4054			TSMPLLPFQRLCVYI
000	2000	A	4874	2	437	FFFLRRSFAFVAQAGVQWCDLGSPQPLPPGF
						K*FSCLSLPSSWDYRHAPPPCPS*FLYF**RQG
						FTMLARLVLNS*PHDLPTSPSQSAEIKGVSHR CPASFYLFLKYYLEAKFCA*GECAPSAGVGA
						GYKRGHKSCLLINCVVQI
651	2001	A	4898	1701	771	DAWGPETRLARILNPDSFIEPRPGRLPELEATR
		1			i	PHMEPKASCPAAAPLMERKFHVLVGVTGSV
. <b>.</b>						AALKLPLLVSKLLDIPGLEVAVVTTERAKHFY
'	l	1		1		SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL
	- 1					RRWADLLLVAPLDANTLGKVASGICDNLLTC
1 1		1	i		ľ	VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ
ļ [		Į				VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP
		I				LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA
						LSLPRGQACSRLNWTEGPGLSFFQPGEAAA
652	2002	A	4927	1	611	FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA
	1		1			SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL
.	İ					PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR
} }	J	j	J	. ]		GQSPIPSRASSPSCSWAQVPGVALARCAGVC
		j	j			KPGDSWRVAACISGRCCSRGRRRGSGPRNPE
		j	i			QSFRGAWGPSFWGSWKSQRELSAGGAQAWP
653	2003	A	4965	2	283	LLGSAGSGLRGEA
""	2003	^	4703	-	203	FFFFI*DGVSLCHPGWNAVARSWLTATSASR
				1		VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTFLAILVLNS*PQVICPPWPPKVLTLQA
654	2004	A	4968	3 .	437	RPGIPGRRFRRSWFCQLP*EPEPGLESLATPGD
[		- I		- '		IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW
	1	l	- 1	İ	·	DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP
	j	J	1		]	PPCLTHLAAASCVVVWCGRWKRDSAECQCD
						HSCSAVSQQEDRCRSSSCS
655	2005	A	4983	201	397	MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN
1	2025			_ <u></u>		TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC
656	2006	A	4988	332	159	LVHKDMYREFFEEEAQASNKHVTRCLTSLVI
657	2007	<del>_</del>	5000			REVHIKTMR*HFLPIRLEKNKNNIKD
021	2007	В	5008	129	465	MAGMKTASGDYIDSSWELRVFVGEEDPEAES

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		l	***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1			İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			İ	peptide	sequence	/=possible nucleotide deletion, \=possible
1			'	sequence	•	nucleotide insertion
<del></del>		<b></b>		sequence		
		]				VILRYTGESHIGGVLLKIVEQINRKQDWSDH
1		1		ł	1	AIWWEQKRQWLLQTHWTLDKYGILADARLF
658	2008	A	5017	1	292	FGPQHRPVILRLPNRRALRLX*
050	2000	^`	3017	1 *	292	FFFFKETESHSVTQAGVQWHDLGSLQPPPPGF
Í					<u> </u>	KRFSCLSLLSSWDYRCAPPHPANFVFLVETGF
659	2009	A	5018	17	338	HHVAQAGLKLLTL*SANLGLSTSLPIPLFILLS
00)	2009	^	3010	17	338	RGHGGKSLTGGTPGNWGDGLLVSEDWSHLIF
	1					T*NSLVSPVLGKWSPCLQGPGLSAVHTWPWL
				i	ļ	MAACWAVHVKTHMRPGLAVLPRLVLNSWS
660	2010	A	5028	2	310	*AIILLWPPKALGLQA
000	2010	Α	3026	2	310	SRVDDFVGERRGGCDECLCGHRGLRAVPLG
			i			HPGHLCLQPPGGPA*FLDYCRGCCPHPVPGST
1			٠ .			AGSCPRQKKTTPGPTVLCVCSFWIYQRGEPH
661	2011	A	5050	752	431	HRTGARWNH
001	2011	Δ	3030	132	431	RQSCSSTQAKVQWFHYGPLQSQPPGLKQSSQ
1 1						LSLPNSRDHRHVPPRLAIFSFAETGSPYFAQAS
] ]						LELLGSSHPPTSASQSARITGVSHRAWPLK*F
662	2012	Α	5054	48	102	NLNQYQTLTMN
002	2012	А	3034	46	103	ELNNGPFQMPLCNGGNLAVTGSWADRSPLH
1						EAASQGRLLALRTLLSQGYNVNAVTLDHVTP
1 1						LHEACLGDHVACARTLLEAGANVNAITIDGV
						TPLFNACSQGSPSCAELLLEYGAQAQLESCLP
						SPTHEGASKGHHECLDILISWGIDVDQEIPHSG
						TPLYVACMAQQFHCIWNLIYAGAGVRKGKY
663	2013	Α	5066	951	580	WDTPLPGAGHQSTQKLE*LFAMVEIWQ
	2015	11	3000	<i>33</i> 1	360	VRNS*SFAHCASVYKHHYMDGQTPCLFVSSK ADLPEGVAVSGPSPAEFCRKHRLPAPVPFSCA
1			}			GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ
664	2014	Α	5071	550	1	LSFIEVLSMEQVNKTVVREFVVLGFSSLARLQ
			5071	330	•	QLLFVIFLLLYLFTLGTNAIISTIVLDRALHTP
						MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK
í I	1	•				TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR
						YMAICNPLRYS VLMGHGVCMGLMAAAWAC
						GFTYSLVTTSLVFHLPFHSSNOHE
665	2015	A	5074	496	692	QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG
				.,,	0,2	PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE
						HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD
	ļ					HDQNEGFHCREECRILGHSDRCWMPRNPMPI
						RSKSPEHVRNIIALSIEATAADVEAYDDCGPT
			.			KRTFATFGKDVSDHPAEERPTLKGKRTVDVT
	- 1					ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL
						PTKPSVSYEIVDPGITARRC
666	2016	A	5080	408	248	IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL
						VGPKAQLSGLQLQPCLYKRR
667	2017	Α	5081	129	247	DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ
	i					PCVY*RR
668	2018	A	5086	852	233	NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF
İ						RALPTIFADIENLKYLLFIRDASOPFYLGHTV
ĺ					ļ	IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE
						TCADOSVIWKLSEDKOLAICLKYAGVHAENA
·	ľ	ŀ			İ	EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG
l l				l		CCSDMAITFNGLTPOKMEVMMYGLYRLRAF
	İ			l		GHYFNDTLVFLPPVGSEND
669	2019	Ā	5101	1	329	PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG
		•	3101	•		ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP
				ļ		RGCQHEAAPCPRGPGSDGLHHASAACASLPP
		-				SPILPVLLPELGPL
670	2020	Ā	5102	3	547	DAWGNRCAVGAAPRLIHLHLCCTPADPSRKP
			3.02		<del></del> /	DAWOUNGATONALKEINENIECTIADISKKI

SEQ ID NO: of nucl- cotide scq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Ghutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \text{\text{\text{\text{-}possible}}} nucleotide insertion  DEL*NMNGRVDYLVTEEENLTRGPSGLGFNI VGGTDQQYVSNDSGIYVSRIKENGAAALDGR LQEGDKILSVNGQDLKNLLHQDAVDLFRNA GYAVSLRVQHRLQVQNGPIGHRGEGDPSGIPI
671	2021	A	5105	672	400	FMVLVPVFALTMVAAWAFMRYRQQL RDGREELCLQQEPTLPSRICSSAPLLYFLFICPF VLLLLLLISLLCLYWKARKLSTLRSNTRKEKA LWVDLKEAGGVTTNRMED*EEDECN
672	2022	A	5148	72	314	IIYFSYNIFLKITELLNDVERLKQALNGLSQLT YTSGNPTKRQSQLIDTLQHQVKSLEQQLAVS NQAHGALQEYVLAPCS
673	2023	Α	5152	210	335	REILCSRIGRINIV*MSLFPNLTCRINAIPIKIPA NHFVEVT
674		A .	5153	3	2953	LTEDQFFDILQKSLQEANTTEQTLAEEAYLDA SIGSSQQFAQAQLHPSSSASFTQASNVSNYSG QTLQPIGVTHVPVGASFASNTVGVQHGFMQH VGISVPSQHLSNSSQISGSGQIQLIGSFGNHPS MMTINNLDGSQIILKGSGQQAPSNVSGGLLV HRQTPNGNSLFGNSSSSPVAQPVTVPFNSTNF QTSLPVHNIIIQRGLAPNSNKVPINIQPKPIQM GQQNTYNNNLGIQQHHVQGISFASASSPQ GSVVGPHMSVNIVNQQNTRKPVTSQAVSSTG GSIVHSPMGQPHAPQSQFLIPTSLSVSSNSVH HVQTINGQLLQTQPSQLISGQVASEHVMLNR NSSNMLRTNQPYTGPMLNNQNTAVHLVSGQ TFAASGSPVIANHASPQLVGGQMPLQQASPT VLHLSPGQSSVSQGRPGFATMPSVTSMSGPSR FPAVSSASTAHPSLGSAVQSGSSGSNFTGDQL TQPNRTPVPVSVSHRLPVSSKSTSTFSNTPGT GTQQQFFCQAQKKCLNQTSPISAPKTTDGLR QAQIPGLLSTTLPGQDSGSKVISASLGTAQPQ QEKVVGSSPGHPAVQVESHSGGQKRPAAKQ LTKGAFILQQLQRDQAHTVTPDKSHFRSLSD AVQRLLSYHVCQGSMPTEEDLRKVDNEFETV ATQLLKRTQAMLNKYRCLLLEDAMRINPPAE MVMIDRMFNQEERASLSRDKRLALVDPEGFQ ADFCCSFKLDKAAHETQFGRSDQHGSKASSS LQPPAKAQGRDRAKTGVTEPMNHDQFHLVP NHIVVSAEGNISKKTECLGRALKFDKVGLVQ YQSTSEEKASRREPLKASQCSPGPEGHRKTSS RSDHGTESKLSSILADSHLEMTCNNSFQDKSL RNSPKNEVLHTDIMKGSGEPQPDLQLTKSLET TFKNILELKKAGRQPQSDPTVSGSVELDFPNF SPMASQBNCLEKFIPDHSEGVVETDSILEAAV NSILEC
675	2025	A	5154		1880	LKKMEPFSCDTFVALPPATVDNRIIFGKNSDR LYDEVQEVVYFPAVVHDNLGERLKCTYIEID QVPETYAVVLSRPAWLWGAEMGANEHGVCI GNEAVWGREEVCDEEALLGMDLVRLGLERA DTAEKALNVIVDLLEKYGQGGNCTEGRMVF SYHNSFLIADRNEAWILETAGKYWAAEKVQE GVRNISNQLSITTKIAREHPDMRNYAKKGW WDGKKEFDFAAAYSYLDTAKMMTSSGRYCE GYKLLNKHKGNITFETMMEILRDKPSGINME GEFLTTASMVFILPQDSSLPCHFFTGTPDPER SVFKPFIFVPHISQLLDTSSPTFELEDLVKKKS HFKPDRRHPLYQKHQQALEVVNNNEEKAKI MLDNMRKLEKELFREMESILQNKHLDVEKIV NLFPQCTKDEIQIYQSNLSVKVSS

NO: of		Met	SEQ	Predicted	Predicted end	Amino scid segmence (A=Alonine ( \= \ \coteine
	SEQID NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ľ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		İ	j	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
į	]	]	J	peptide	30423300	/-possible nucleotide deletion, \-possible
		İ		sequence		nucleotide insertion
676	2026	A	5155	2	306	FFFLRRSLALSPRPDCGLQWRNLGSLQAPPPG
				-	500	FTPFSCLSLPSSWDYRRPPPRPANFLYF**RRG
	l	l	1		1	FTLLARMVSIS*PHDPPASASQSAGITGVSHRA
				ļ	1	RPT
677	2027	A	5167	97	740	FFHSVDLLALEQSKTFYKPDWFDIVESEVKCC
			1			KEAVCVIDMSSFTEFEITSTGDQALEVLQYLF
	1		i			SNDLDVPVGHIVHTGMLNEGGGYENDCSIAR
				}		LNKRSFFMISPTDQQVHCWAWLKKHMPKDS
			1			NLLLEDVTWKYTALNLIGPRAVDVLSELSYA
·						PMTPDHFPSLFCKEMSVGYANGIRVMSMTHT
						GEPGFMLYIPIEYRWGFTMLSTLVSNS
678	2028	Α	5183	1919	2018	PALCRLRDDMTVCVADFGLSKKIYSGDYYRO
						GRIAKMPVKWIAIESLADRVYTSKSDVWAFG
	·					VTMWEIATRGMTPYPGVQNHEMYDYLLHG
Ì	•					HRLKQPEDCLDELCKI**SPQSP
679	2029	A	5190	39	499	RESQVKHFKMRKIDLCLSSEGSEVILATSSDE
					""	KHPPENIIDGNPETFWTTTGMFPQEFIICFHKH
		'				VRIERLVIQSYFVQTLKIEKSTSKEPVDFEQWI
			ł			EKDLVHTEGQLQNEEIVAHDGSATYLRFIIVS
						AFDHFASVHSVSAEGTVVSNLSS
680	2030	Α	5204	541	92	EILAVLKLACGDISLNALALMVATAVLTLAPL
					-	LLICLSYLFILSAILRVPSAAGRCKAFSTCSAH
ĺ						RTVVVVFYGTISFMYFKPKAKDPNVDKTVAL
						FYGVVTPSLNPIIYSLRNAEVKAAVLTLLRGG
ļ	ſ					LLSRKASHCYCCPLPLSAGIG
681	2031	A	5207	10	247	VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI
1						KNESPLPGTLAHTCNTSTLGGRGRWIT*GREF
						DTSMANMVKPCLYRK
682	2032	A	5210	2	231	FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F
1						SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC
						WPGWSRTPDLS
683	2033	Α	5218	85	402	CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF
	ł					MIILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG
Į.						YARCYTFLIF*LFL*IPADSPTDQEPKTVMLSK
					<b>,</b>	QSESAI
684	2034	A	5220	1	194	NLMKEMQNLNSENHKTWEEYKDTK*IMSYF
	ı					YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL
	l					TDS
685	2035	A	5228	260	440	LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED
<u> </u>						QKKTPQVPINFTELSKCS*S*KIMSGERE
686	2036	Α	5239	79	508	GGEAAARAAKLSSPRPHRVGRRERGVGGMS
. [	ſ	ĺ		ĺ	i	AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR
		l		·		KHSRPIVTVWERELRKAKPNRKLTFLYLAND
						VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD
				<u> </u>	· · _	ESCKKHLGRVLSIWEERS
687	2037	A	5244	1	428	MAAVVAATALKGRGARNARVLRGILAGATA
1		ĺ	1		ľ	NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI
1		į				PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR
	- 1	l				EVDKDRVKQMKARQNMRLSNTGEYESQRFR
	}					ASSQSAPSPDVGSGVQT
688	2038	A	5249	1	1407	LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP
		- 1		[	ĺ	SGDSDLATALHRLSLRRQNYLSEKOFFAEEW
- 1	İ	l		1	l	QRKIQVLADQKEGVSGCVTPTESLASLCTTQS
- 1		i		1	l	EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY
İ		i		- 1	ł	HWQQLAQPNLGTILDPRPGVITKGFTQLPGD
J	1	l	,	1	į	AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS
		ſ	{	1	ĺ	KPVTGIFLPPITSAGGPVTVATANPGKCLSCT
			į			NOTE THE PROPERTY OF THE PROPE
.	ŀ	ı	1	ı	i	NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS

SEQ I	SEQID	Met	SEQ	Predicted	Predicted end	L Amiros and a second a second and a second and a second
NO: of		hod	ID NO:		nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide		in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ſ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		Í	Í	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	1	1	peptide	_	/=possible nucleotide deletion, \-possible
				sequence		nucleotide insertion
1						FSSTMSLAKLLQERGISAKVYHSPISENPLQPL
		ĺ	[			PKSLAIPSTPPNSPSHSPCPSPLPFEPRVHI.SEN
1	1			i	1	FLASRPAETFLQEMYGLRPSRNPPDVGOLKM
1			1	l	ļ	NLVDRLKRLGIARVVKNPGAOENGRCOEAEI
	1		1			GPQKPDSAVYLNSGSSLLGGLRRNOSLPVIM
689	2039	A	5254	<del> </del>	260	GSFAAPVCTSSPKMGVLKED
1 005	2039	1^	3234	2	2621	LSLFGSRALGRSGARAMAKAKKVGARRKAS
1	1	1		ł		GAPAGARGGPAKANSNPFEVKVNRQKFQILG
	1		]	]		RKTRHDVGLPGVSRARALRKRTQTLLKEYKE
		1			ſ	RDKSNVFRDKRFGEYNSNMSPEEKMMKRFA
1	1	1	1	ĺ	l	LEQQRHHEKKSIYNLNEDEELTHYGQSLADIE
			1			KHNDIVDSDSDAEDRGTLSGELTAAHFGGGG
1	j	J		ļ	J	GLLHKKTQQEGEEREKPKSRKELIEELIAKSK
1		1		1	İ	QEKRERQAQREDALELTEKLDQDWKEIQTLL SHKTPKSENRDKKEKPKPDAYDMMVRELGF
1	1	İ	j	ĺ	ĺ	EMKAQPSNRMKTEAELAKEEQEHLRKLEAE
i	1	1	ł	ļ		RLRRMLGKDEDENVKKPKHMSADDLNDGFV
ļ	1	1		J .		LDKDDRRLLSYKDGKMNVEEDVQEEQSKEA
1	ľ	i				SDPESNEEEGDSSGGEDTEESDSPDSHLDLES
[	1	1				NVESEEENEKPAKEQRQTPGKGLISGKERAG
.1	i		1			KATRDELPYTFAAPESYEELRSLLLGRSMEEQ
	i	1				LLVVERIQKCNHPSLAEGNKAKLEKLFGFLLE
1		]				YVGDLATDDPPDLTVIDKLVVHLYHLCOMFP
	1	1	[			ESASDAIKFVLRDAMHEMEEMIETKGRAALP
		ĺ	İ			GLDVLIYLKITGLLFPTSDFWHPVVTPALVCI
1		1				SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS
1	j			ļ		QRFIPELINFLLGILYIATPNKASQGSTLVHPFR
}	1					ALGKNSELLVVSAREDVATWQQSSLSLRWA
ĺ	.[	[				SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM
ł	i	l i			-	YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ
]		l l				ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL
				i		FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE
<u></u>	1	f i		i		RKRKVKQLFNSLATQEGEWKALKRKKFKK
690	2040	A	5261	1	304	FFFFVFLVETGFHHVGQAGLELLTSGDPPTW
l	1			}		ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL
ļ		J	J			FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT
						SFVK
691	2041	Α	5270	3	158	NCHTTHCTANWVHLPGTPPGWKIDGPAAAL
<u></u>	10015					EVLSSFFFFFLKFSYKPQNIV
692	2042	Α	5282	56	1268	GMEPVGCCGECRGSSVDPRSTFVLSNLARVV
					ļ	ERVLTFLPAKALLRVACVCRLWRECVRRVLR
	[ [	ľ	1		í	THRSVTWISAGLAEAGHLEGHCLVRVVAPEL
		ŀł	ł	ł		ENVRILPHTVLYMADSETFISLEECRGHKRAR
		I		1	J	KRTSMETALALEKLFPKQCOVLGIVTPGIVVT
	] 1			1		PMGSGSNRPQEIEIGESGFALLFPOIEGIKIOPF
	j l	1	ſ	1	ľ	HFIKDPKNLTLERHOLTEVGLLDNPELRVVLV
	l 1	- 1	i	ł.		FGYNCCKVGASNYLQQVVSTFSDMNIILAGG
		1				QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI
	] }	]			Ì	QSATVLLNEDVSDEKTAEAAMQRLKAANIPE
	] [		1	ĺ	1	HNTIGFMFACVGRGFQYYRAKGNVEADAFR
I	!	- 1		ł	1	KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN
693	2043	A	5301	362	507	EVKDDDLFHSYTTIMALIHLGSSK
				- VZ	-01	EEIKERFGPGLVIYWYGFIQELDCNRERGILLK
694	2044	A	5310	1	204	ACFPTNIVTLCHSIA
		1		- 1	~~~	RVLTAINHTLKENLRKFYKGKKDKPLDLRPK
		- 1		1		KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA
695	2045	A	5315	125		ETRSTAVKSEVQVCISLLLCLEDRTMPKKAKP
						Z-1-W-11-1-WDD-YQ-YCIDLLLUCLEUKIMPKKAKP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	İ	USSN	location		
Seq-	uence	1	09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
uence	ucuicc		914			M=Methionine, N=Asparagine, P=Proline,
ucuce			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ŀ			İ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ł		İ		peptide		/-possible nucleotide deletion, \-possible
<u></u>				sequence		nucleotide insertion
				1		TGSGKEEGPAPCKQMKLEAAGGPSALNFDSP
		l				SSLFESLISPIKTETFFKEFWEQKPLLIQRDDPA
1						LATYYGSLFKLTDLKSLCSRGMYYGRDVNV
1			ļ			CRCVNGKKKVLNKDGKAHFLQLRKDFDQKR
			•		ł	ATIQFHQPQRFKDELWRIQEKLECYFGSLVGS
			l		ŀ	NVYITPAGSQGLPPHYDDVEVFILQLEGEKH
)		1	}	ļ	J	WRLYHPTVPLAREYSVEAEERIGRPVHEFML
}					Į.	KPGDLLYFPRGTIHQADTPAGLAHSTHVTIST
1						YQNNSWGDFLLDTISGLVFDTAKEDVELRTG
						IPRQLLLQVESTTVATRRLSGFLRTLADRLEG
						TKELLSSDMKKDFIMHRLPPYSAGDGAELSTP
						GGKLPRLDSVVRLQFKDHIVLTVLPDQDQSD
1						ETQEKMVYIYHSLKNSRETHMMGNEEETEFH
						GLRFPLSHLDALKQIWNSPAISVKDLKLTTDE
						EKESLVLSLWTECLIOVV
696	2046	A	5318	1476	742	
"	2010	Λ.	3310	1470	742	LMKXYLEAAELGEISDIHTKLLRLSSSQGTIET
1						SLQDIDSRLSPGGSLADAWAHQEGTHPKDRN
1 . 1						VEKLQVLLNCMTEIYYQFKKDKAERRLAYN
						EEQIHKFDKQKLYYHATKAMTHFTDECVKK
1 1						YEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI
						EEEVSKYQEYTNELQETLPQKMFTASSGIKHT
1						MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE
697	2047		5300	0.7	.50	LAENNHILESGGSLTMDGGLRNVDCL
09/	2047	A	5320	244	478	LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP
1 1						PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG
						VSPSWPGWSRTPDFR
698	2048	Α	5324	266	714	LPIRKSLRSVRSGFPTSQSPITRNLDGTASGSC
]	i					LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP
1 1						VGGLLMAFQKYSGETVQERKQKDRKALHEL
						KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA
						KKIEALLNLPRNPSVIDKQDKD
699	2049	A	5334	699	277	RPHGHLVCISSSAGLSGVNGLADYCASKFAA
1 1	- 1				·	FGFAESVFVETFVQKQKGIKTTTVCPFFIKTGM
i i	1					FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM
1 1	l	ļ				YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI
L I	ı	1				LHAMDGFADQKK
700	2050	A	5344	3	614	PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE
				_	···	QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS
ł l						QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG
						PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV
<u> </u>				J	<b> </b>	VETIQAQLLSTHDQPSVQALADEKNGAQTRP
j l				J	•	AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK
					-	LLATNGIPL
701	2051	A	5346	3	1383	
l l	2001		JJ-10	·		HASVLFCRVMAASKTQGAVARMQEDRDGSC
i I		Ì		1		STVGGVGYGDSKDCILEPLSLPESPGGTTTLE
	I	i		į	ļ	GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV
!	- 1					IADVKLVADFQRYILYWRKRFTEQPITDFCSV
] ]		1				IRINSTAPFEEQENYFLLCDVLPEDRILREELQ
1 1			)	1	1	KQRLREILEQQQQERNDTNFHGVCMFCNEEF
1	İ	ŀ		ł	1	LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL
		l		1	ĺ	CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR
	ı			ŀ	İ	KKQHRKINPKNREYDRFYVINYLELGKSWEE
						VQLEDDRELLDHQEDDWSDWEEHPASAVCL
	İ	ŀ		i		FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG
				I	•	LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS
				i		KADLRTHMEETKHTSLLPDRKTWDQLEYYFP
[	ſ	[	-	i	1	TYENDTLLWTLSDSESDLTAQEQNENVPIISE
<u> </u>		1	1			DTSKLYALKQSSILNQLLL
702	2052	A	5356	2502	1540	MAAATRGCRPWGSLLGLLGLVSAAAAAWD
						The state of the s

SEQID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine O=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isolencine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LASLRCTLGAFCECDFRPDLPGLECDLAQHL AGQHLAKALVVKALKAFVRDPAPTKPLVLSL HGWTGTGKSYVSSLLAHYLFQGGLRSPRVH HPSPVLHFPHPSHIERYKKDLKSWVQGNLTA CGRSLFLFDEMDKMPPGLMEVLRPFLGSSWV VYGTNYRKAIFIFISNTGGEQINQVALEAWRS RRDREEILJQELEPVISRAVLDNPHHGFSNSGI MEERLLDAVVPFLPLQRHHVRHCVLNELAQL GLEPRDEVVQAVLDSTTFFPEDEQLFSSNGCK TVASRIAFFL
703	2053		5380	278	657	LFLQKLRMKTEEEARTHTEIEMFLRKEQQKL EERLEFWMEKYDKDTEMKQNELNALKATKA SDLAHLQDLAKMIREYEQVIIEDRIEKERSKK KVKQDLLELKSVIKLQAWWRGTMIRREIGGF KM
704	2054	A	5381		1003	FRGRAVKMAAVVEVEVGGGAAGERELDEV DMSDLSPEEQWRVEHARMHAKHRGHEAMH AEMVLILIATLVVAQLLLVQWKQRHPRSYN MVTLFQMWVVPLYPTVKLHWWRFLVIWILF SAVTAFVTFRATRKPLVQTTPRLVYKWFLLIY KISYATGIVGYMAVMFTLFGLNLLFKIKPEDA MDFGISLLFYGLYYGVLERDFAEMCADYMA STIGFYSESGMPTKHLSDSVCAVCGQQIFVDV SEEGIIENTYRLSCNHVFHEFCIRGWCIVGKK QTCPYCKEKVDLKRMFSNPWERPHVMYGQL LDWLRYLVAWQPVIIGVVQGINYILGLE
705	2055	A	5396	3	675	IYDRDPLQLATRAGQPLDINMAGEPKPYRPKP GNKRPLSALYRLESKEPFLSVGGYVFDYDYY RDDFYNRLFDYHGRVPPPPRAVIPLKRPRVA VTTTRRGKGVFSMKGGSRSTASGSTGSKLKS DELQTIKKELTQIKTKIDSVLGRLDKIEKQQK AEAEAQKKLLEESLVLIQEECVSEIADHSTEEP AEGGPDADGEEMTDGIEEAFDEDGGHELFLQ IK
706	2056	A	5410	2	98	GRVGLNLEGRGCSEPKWRHCTPTWATEQDSI
707	2057	A	5415	6	287	PFKLTPSFLSHAFSSGQERKVFIELNHIKKCNT VRGVFVLEEFGNYTILLLGLDSHGSNSNLGAP EEGLGAGRKRTSVEKSGGAGVTRKKRDP
708	2058	A	5423	3	291	SSSNPLGSPSTLWKLCSFVLHNKSCCCSFFGS TPTLRAITLTVRVCGFIPEVSKTTNPLGRTNNS GCTIFKTVTLTARSTASLLKSVRPRTHQKE
. 709	2059	A	5424	679	347	RIRHEEKRGSRGRGRRTSEEDTPKKKKHKGG SEFTDTILSVHPSDVLDMPVDPNEPTYCLCHQ VSYGEMIGCDNPDCPIEWFHFACVDLTTKPK GKWFCPRCVQEKRKKK
710	2060	A	5442	1073	559	QESLKKKIQPKLSLTLSSSVSRGNVSTPPRHSS GSLTPPVTPPITPSSSFRSSTPTGSEYDEEEVDY EESDSDESWTTESAISSEAILSSMCMNGGEEK PFACPVPGCKKRYKNVNGIKYHAKNGHRTQI RVRKPFKCRCGKSYKTAQGLRHHTINFHPPV SAEIIRKMQQ
711	2061	A	5449	1	319	GDSLCVPQYNKYREERVILFLKMASGHAFQP DLVKRIRDAIRMGLSARHVPSLILETKGIPYTL NGKKVEVAVKQIIAGKAVEQGGAFSNPETLD LYRDIPELQGF
712	2062	A	5499	91	749 ·	RPTPGHGDFWMQPLTKDAGMSLSSVTLASAL QVRGEALSEEEIWSLLFLAAEQLLEDLRNDSS DYVVCPWSALLSAAGSLSFQGRVSHIEAAPF

NO: of No: of Porphile coulded and continuous properties of the properties of the properties of the properties of the properties of properti	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
tence colde to correspond to correspond to correspond to the tence of the corresponding to th			hod				
	nucl-	peptide		in	, , ,		
1914   1914	eotide	1	1	USSN			
Part	seq-		ĺ			to last amino	
minio seid residue of sequence   residue				1			O=Chitamine R=Arrining S=Sering
residue of peptide sequence   Y=Tyrosine, X=Usknown, =Sinp codon,   Popsible nucleotide cleti	1		·	[			T=Threenine V=Veline W=Truntonhon
Peptide	ì						V=Turcine V=Introum *-Ctom codes
						Sequence	Enoscible prolectide deletion \ socible
			ļ				purioside insertion
	<del></del>	<del> </del>	<del> </del>	<del>                                     </del>	Saquata	<del></del>	
RRCTLQSVLEACRVHEKEVSYYPAPAGLET    713   2063   A   5506   22   478   VEELLUSRIDPHLITPMYSFLAHLSED_ESSESSESCENCVAIK    VEELLUSRIDPHLITPMYSFLAHLSED_ESSESSESCENCVAIK    VEELLUSRIDPHLITPMYSFLAHLSED_ESSESSESSESSESSESSESSESSESSESSESSESSESS	į.	}	l	Ì	ļ	•	
713				1			PROTE OCCUR A CONTRICTOR A DA OL VED
1713							RRC1LQSVLEACKVHEKEVSVYPAPAGLHIK
TSSIPQLLYNLNCCDKTISYMGCAIQLFLETGL   GOVECLLAVMAYDRCVAICKRLETYMYNIPR   GGWECLLAVMAYDRCVAICKRLETYMYNIPR   FILCRIL AVMAYDRCVAICKRLETYMYNIPR   FILCRIL AVMAYDRCVAICKRLETYMYNIPR   FILCRIL AVMAYDRCVAICKRLETYMYNIPR   FILCRIL AVMAYDRCVAICKRLETYMYNIPR   FILCRIL AVMAYDRCVAICKRLETYMYNIPR   FILCRIL AVMAYDRCVAICKRLETYMYNIPR   FILCRIL AVMAYDRCVAICKRLETYM   FILCRIL AVMAYDRCVAIC	713	2063		5506	22	470	KLVGLVLGTISEVSKEPCFSSSSCWSCVAIKI
	/13	2000	^	٥٥٥٥	22	470	TESTION AND NOOPERATE AND THE PER
PRICRGIA SYTYWGCGYANSLAMSPYTIRIPR   CGHEVDHIFLEMPALIRMAGISTY						ŀ	
114	1			<b>i</b>			
1714	1	ľ					PRICKGLYSVIWGCGVANSLAMSPYTLRLPR
	714	2064	_	6614	25		
DK	/14	2004	A	2214	لک	220	
715	ł	·				f .	
GRTALPHHSGGSSGYESIRRDSEATGSASSAP	715	2066		5506		-	
DSMSESGAASFOARTRSLKSPKKRATGLQRR	'15	2000	A.	3326	3	810	LVIAPRRPQRYSSGHGSDNSSVLSGELPPAM
RILPAPLPDITALGRKPSLPGQWYDLPPPLAG   SLKEPFEIK VYEDDUGLQRPRTPREAPTQG  LACVSTRLRLAERRQQRLREVQAKIKHLCEE  LAETQGRLMLEPGRWLEQFEVDPELEPESARE   VIAALERTAALEQCVNLCKAHVMWYTCFD  ISVAASAAIPGPQEVDV    716   2066   A   5529   458   790   SPGYGENFTVTSXNIAVPLCEMNKIYSYYSD    SSSSERTMDLVLEMCNTNSIHWCGISGRQLG  KLHPSSLCLALILLSSVQGLQSISGLRLTDTF  LKRTYEYDDIAQVCV    TIT   2067   A   5531   3   460   NSEDLLKYFNPESWQEDLDNWYLDTPRYRG    RSYHDDKSKVDLDRLNDAKRYSCTPRNYS  VNIREELKLANVVFFPRCLLVQRCGGNCGCG  TVNWRSCTCNSGKTVKKYBLEVLQFEPGHIKKY    TVNWRSCTCNSGKTVKKYBLEVLQFEPGHIKK    TVNWRSCTCNSGKTVKKYBLEVLGFEPGHIKK    TVNWRSCTCNSGKTVKKYBLEVLGFEPGHIKK    TVNWRSCTCNSGKTVKKYBLEVLGFEPGHIKK    TVNWRSCTCNSGKTVKKYBLEVLGFEPGHIKK    TVNWRSCTCNSGKTVKKYBLEVLGFEPGHIKK    TVNWRSCTCNSGKTVKKYBLEVLGFEPGHIKK    TVNWRSCTCNSGKTVKKYBLEVLGFEPGHIKK    TVNWRSCTCNSGKTVKKYBLEVLGFEPGHIKK    TVNWRSCTCNSGKTVTACKNSBLOQRVCVCPC    HGMACTARGECCHTECLGGCSQPEDFRACV    ACRILYPQGACLWACPPGTVQTSSWCCYTAC    ACRILYPQGACLWACPPGTVQTSSWCCYTAC    TVNWRSCTNSGKTVTACKNSBLOQRVCCPC    HGMACTARGECCHTECLGGCSQPEDFRACV    ACRILYPQGACLWACPPGTVQTSSWCCYTAC    ACRILYPQGACLWACPPGTVQTSSWCCYTAC    ACRILYPQGACLWACPGTTQVCSCACACPCPCPC    TVNTON TOTAL TVNTON TOTAL TOTAL TVNTON TOTAL TVNTON TOTAL TVNT							GKIALFHHSGGSSGYESLRRDSEATGSASSAP
SIKEPFEIKVYEIDDVERI.QRPRPTPREAPTOGE	ł i						
LACVSTRIKIAERRQQRIMEPOFEVPDELPPESAE							RLIPAPLPDTTALGRKPSLPGQWVDLPPPLAG
LAETQGRILMIEPGRWILEGPENDFELEPESAE					i		SLKEPFEIKVYEIDDVERLQRPRPTPREAPTQG
YIAALERATAALEQCVNICKAHVMMVTCFD   ISVAASAABCOPQDV   ISVASAABCOPQDV   ISVASAABCOPQEDV   ISVASAABCOPQEDV   ISVASAABCOPQEDV   ISVASAABCOPQEDV   ISVASAABCOPQEDV   ISVASAABCOPQEDV   ISVASAABCOPQEDV   ISVASAABCOPQEDV   ISVASAABCOPQEDV   ISVASAABCOPQEDV   ISVASCILARITLISSVQGLQSISQGLRITDTF   LKRTYEYDDIAQVCV   LKRTYEYDDIAQVCV   ISVASCILARITLISSVQGLQSISQGRITDTF   LKRTYEYDDIAQVCV   ISVASCILARITLISSVQGLQSISQGRITDTF   LKRTYEYDDIAQVCV   ISVASCILARITLISSVQGLQSISQGRITDTF   LKRTYEYDDIAQVCV   ISVASCILARITVENED   ISVAS							LACVSTRLRLAERRQQRLREVQAKHKHLCEE
ISVAASAAPGPQEVDV	1						LAETQGRLMLEPGRWLEQFEVDPELEPESAE
716	1 1						YLAALERATAALEQCVNLCKAHVMMVTCFD
SSSSERTMDLVLEMCNTNSHWCGISGRQLG   KLHPSSLCLAITLLSSVQGLQSISGLRITDTF   LKRYPEYDDIAQVCV     717   2067	L						ISVAASAAIPGPQEVDV
	716	2066	A	5529	458	790	
LKRTYEYDDIAQVCV							
717							KLHPSSSLCLALTLLSSVQGLQSISGLRLTDTF
RSYHPRKSKVDLDRLNDDAKRYSCTPRNYS	L						LKRTYEYDDIAQVCV '
RSYHDRKSK VDLDRLNDDAKRYSCTPRNYS VNIRBELKLANVVFFPRCLLVQRCGGNCGCG TVNRRSCTCNSGKTVKKYHEVLQFEPGHIKR RGRAKTMALVDIQLDHHERCDCICSSRPPR  718 2068 A 5586 311 88 AVLKNMAPMTALGLLDHILNILIFLSAGEDF TSVV SEIMMYILLVFLTUMLIEMIYCYPKVS KAEEAAQENA  719 2069 A 5598 1 330 KNCANEAVQKILDRVLSRYDVRLRPNFGSM LATNSTGLNEDELMAHQQEKDSSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FFLAFGLFNIVAAERC  720 2070 A 5628 798 148 LPPAQPEAULIANVVVVIILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVVAGV LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWLAGRYERASQGPASHSRFSRDRG  721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLIFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GILQPAPGANIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHIDYRCWTSSHOQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYPQGACLWACPPGTYQYESWRCVTA ERCASLISVPGRASTFGIHQGSCLAQCPSGFT	717	2067	A ]	5531	3	460	NSEDLLKYFNPESWQEDLDNMYLDTPRYRG
TVNWRSCTCNSGKTVKKYHEVLQFEPGHIKR   RGRAKTMALVTQLDHHERCDCICSSRPPR   RGRAKTMALVTQLDHHERCDCICSSRPPR   A   S586   311   88		1	1			j	
RGRAKTMALVDIQLDHIERCDCICSSRPPR					-	'1	VNIREELKLANVVFFPRCLLVQRCGGNCGCG
RGRAKTMALVDIQLDHIERCDCICSSRPPR	1				ļ	ļ	TVNWRSCTCNSGKTVKKYHEVLQFEPGHIKR
TSVVSEIMMYILLVRLTLWLLIEMIYCYRKVS							
KAEEAAQENA	718	2068	A	5586	311	88	
KAEEAAQENA	-		- 1				TSVVSEIMMYILLVFLTLWLLIEMIYCYRKVS
LATNSTRGLNEDELMAHGQEKDSSSESEDSC PPSPGCSFTEGPSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC  12070 A 5628 798 148 LPFAQIFEAWILLANVVVVLILVPLKDRLIDP LLLRCKILPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYILLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG  121 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELFTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  122 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMP TATGEDFRGLSFPRLTQVITDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	LI						
LATNSTRGLNEDELMAHGQEKDSSSESEDSC   PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL   FPLAFGLFNIVAAERC   PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL   FPLAFGLFNIVAAERC   PPAQPEAWLILANVVVVIILVPLKDRLIDP   LLRCKLLPSALQKMALGMFFGFTSVIVAGV   LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW   WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG   AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH   CPPKDFGNINNCRMDLYFFLLAGIQAVTALLF   VWIAGRYERASQGPASHSRFSRDRG     721   2071   A 5632   146   536   MSALIVRKLRSAELTLFSELPTVLGANVNAA   KLHETALHHAAKVKNVDLIEMLIEFIGGNIYA   RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT   LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID   YLSYN     722   2072   A 5638   3   3806   CPSLDIRSEVAELRQLENCSVVEGHLQILMF   TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES   LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD   VALPALGAVLRGAVRVEKNQELCHLSTIDW   GLLQPAPGANHIVGNKLGEECADVCPGVLGA   AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP   HGMACTARGECCHTECLGGCSQPEDPRACV   ACRHLYPQGACLWACPPGTYQYESWRCVTA   ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	719	2069	A	5598	1	330	KNCANEAVVOKILDRVLSRYDVRLRPNFGSM
PPSPGCSFTEGPSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC  720 2070 A 5628 798 148 LPPAQIPEAWLLLANVVVVLILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGPTSVIVAGV LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSFSRDRG  721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVIDYLLLFRVYGLES LRDLFPNLAVIRGTRLFIGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GILQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYPQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	1 1	ł	ł	- 1			LATNSTRGLNEDELMAHGOEKDSSSESEDSC
720 2070 A 5628 798 148 LPPAQIPEAWILLANVVVVIILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG  721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVIDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	i i	ł	Į			ł	PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFI.
2070   A   5628   798   148   LPPAQIPEAWLLLANVVVVLILVPLKDRLIDP	<u> </u>						FPLAFGLFNIVAAERC
LLLRCKLLPSALQKMALGMFFGPTSVIVAGV LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG  721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  722 2072 A 5638 3 3806 CPSLDIRSEVAPLRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYPQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	720	2070	A	5628	798	148	
LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTAILF VWIAGRYERASQGPASHSRFSRDRG  721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT		ŀ	ļ		j		
WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTAILF VWIAGRYERASQGPASHSRFSRDRG  721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT					1		
AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG  721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFFNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT			.		1	1	WOIPOYLLIGISEIFASIPGLEFAYSEAPRSMOG
CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG  721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT			ı	ŀ	1	ì	
721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT		ļ	- 1	j	1	I	
721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVIDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GILQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT			- 1	ļ	ŀ	. [	
KLHETALHHAAKVKNVDLEMLEFGGNTYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GILQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	721	2071	A	5632	146	536	
RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVIDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GILQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	ļ	ļ		·	· -		
LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN  722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVIDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GILQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT		ŀ		ļ		Ì	
722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	ľ	I		- 1		l	
722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	}	ł	ł	i	ŀ	J	
TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	722	2072	$\frac{1}{A}$	5638	<del></del>	3806	
LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT				2000	·	3000	
VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT			1	j	1	ł	
GLLQPAPGANHIVGNKLGEEČADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT		J			. [	j	
AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT				j	. [	ļ	
HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT				1		İ	
ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	1	ļ		[			
ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT		1	•	ľ		į	
		į	1		į	ŀ	
RNSSSIFCHKCEGLCPKECKVGTKTIDSIOAA	Į	1	- 1	- 1	j		
						<u></u>	KNSSSIFCHKCEGLCPKECKVGTKTIDSIQAA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						QDLVGCTHVEGSLILNLRQGYNLEPQLQHSL GLVETITGFLKIKHSFALVSLGFFKNLKLIRGD AMVDGNYTLYVLDNQNLQQLGSWVAAGLTI PVGKIYFAFNPRLCLEHIYRLEEVTGTRGRQN KAEINPRTNGDRAACQTRTLRFVSNVTEADRI LLRWERYEPLEARDLLSFIVYYKESPFQNATE HVGPDACGTQSWNLLDVELPLSRTQEPGVTL ASLKPWTQYAVFVRAITLTTEEDSPHQGAQS- PIVYLRTLPAAPTVPQDVISTSNSSSHLLVRW KPPTQRNGNLTYYLVLWQRLAEDGDLYLND YCHRGLRLPTSNNDPRFDGEDGDPEAEMESD CCPCQHPPPGQVLPPLEAQEASFQKKFENFLH NAITIPISPWKVTSINKSPQRDSGRHRRAAGPL RLGGNSSDFEIQEDKVPRERAVLSGLRHFTEY RIDIHACNHAAHTVGCSAATFVFARTMPHRE ADGIPGKVAWEASSKNSVLLRWLEPPDPNGL ILKYEIKYRRLGEEATVLCVSRLRYAKFGGV HLALLPPGNYSARVRATSLAGNGSWTDSVAF YILGPEEDAGGLHVLLTATPVGLTILIVLAA
723	2073	A	5672	1	216	LGFFYGKKRNRTLYASVNPEYFSASDMYVPD EWEVPREQISIRELGQGSFGMVYEGLARGLE AGEESTPVALKTVNELASPRECIEFLKEASVM KAFKCHHVVRLLGVVSQGQPTLVIMELMTR GDLKSHLRSLRPEAENNPGLPQPALGEMIQM AGEIADGMAYLAANKFVHRDLAARNCMVSQ DFTVKIGDFGMTRDVYETDYYRKGGKGLLP VRWMAPESLKDGIFTTHSDVWSFGVVLWEIV TLAEQPYQGLSNEQVLKFVMDGGVLEELEGC PLQLQELMSRCWQPNPRLRPSFTHILDSIQEEL RPSFRLLSFYYSPECRGARGSLPTTDAEPDSSP TPRDCSPQNGGPGH LAWIDNILPEKEKKETDKKRKKKGAHEDCD
724	2074	A	5704	4235	940	EEPOFPPPSVIKIPMESVQSDPQNGIHCIARKR SSSWSYSL ARGRRSRPVWAASWGGRGRPAARRPRGLA ATMGFELDRFDGDVDPDLKCALCHKVLEDP
						LTTPCGHVFCAGCVLPWVVQEGSCPARCRGR LSAKELNHVLPLKRLILKLDIKCAYATRGCGR VVKLQQLPEHLERCDFAPARCRHAGCGQVLL RRDVEAHMRDACDARPVGRCQEGCGLPLTH GEQRAGGHCCARALRAHNGALQARLGALHK ALKEALRAGKREKSLVAQLAAAQLELQMT ALRYQKKFTEYSARLDSLSRCVAAPPGGKGE ETKSLTLVLHRDSGSLGFNIGGRPSVDNHDG SSSEGIFVSKIVDSGPAAKEGGLQIHDRIEVN GRDLSRATHDQAVEAFKTAKEPIVVQVLRRT PRTKMFTPPSESQLVDTGTQTDITFEHIMALT KMSSPSPPVLDPYLLPEEHPSAHEYYDPNDYI GDIHQEMDREELELEEVDLYRMNSQDKLGLT VCYRTDDEDDIGIYISEIDPNSIAAKDGRREG DRIIQINGIEVQNREEAVALLTSEENKNFSLLI ARAELQLDEGWMDDDRNDFLDDLHMDMLE EQHHQAMQFTASVLQQKKHDEDGGTTDTAT ILSNQHEKDSGVGRTDESTRNDESSEQENNG DDATASSNPLAGQRKLTCSQDTLGSGDLPFS NESFISADCTDADYLGIPVDECERFRELLELK CQVKSATPYGLYYPSGPLDAGKSDPESVDKE LELLNEELRSIELECLSIVRAHKMQQLKEQYR ESWMLHNSGFRNYNTSIDVRRHELSDITELPE KSDKDSSSAYNTGESCRSTPLTLEISPDNSLRR

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, B=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	l	l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
İ		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ			peptide		/=possible nucleotide deletion, \=possible
	<del> </del>	<del> </del>	<del></del>	sequence	<del></del>	nucleotide insertion
1 .				1	]	AAEGISCPSSEGAVGTTEAYGPASKNLLSITE DPEVGTPTYSPSLKELDPNQPLESKERRASDG
		1		Ī	}	SRSPTPSQKLGSAYLPSYHHSPYKHAHIPAHA
			İ			QHYQSYMQLIQQKSAVEYAQSQMSLVSMCK
1		ŀ				DLSSPTPSEPRMEWKVKIRSDGTRYTTKRPVR
		1	<b>!</b>			DRLLRERALKIREERSGMTTDDDAVSEMKM
						GRYWSKEERKQHLVKAKEQRRRREFMMQSR
						LDCLKEQQAADDRKEMNILELSHKKMMKKR
						NKKIFDNWMTIQELLTHGTKSPDGTRVYNSF
725	2075	A	5707	3	1770	LSVTTV
'	20.3	1	3.0.	•	1770	QISTEVSEAPVANDKPKTLVVKVQKKAADLP DRDTWKGRFDFLMSCVGYAIGLGNVWRFPY
						LCGKNGGGAFLIPYFLTLIFAGVPLFLLECSLG
1						QYTSIGGLGVWKLAPMFKGVGLAAAVLSFW
						LNIYYIVIISWAIYYLYNSFITTLPWKQCDNP
i						WNTDRCFSNYSMVNTTNMTSAVVEFWERN
						MHQMTDGLDKPGQIRWPLAITLAIAWILVYF
						CIWKGVGWTGKVVYFSATYPYIMLIILFFRGV
						TLPGAKEGILFYITPNFRKLSDSEVWLDAATQ IPFSYGLGLGSLIALGSYNSFHNNVYRDSIIVC
						CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV
ŀ						AASGPGLAFLAYPEAVTQLPISPLWAILFFSM
						LLMLGIDSQFCTVEGFITALVDEYPRLLRNRR
i		1				ELFIAAVCIISYLIGLSNITQGGIYVFKLFDYYS
						ASGMSLLFLVFFECVSISWFYGVNRFYDNIQE
						MVGSRPCIWWKLCWSFFTPIIVAGVFIFSAVQ
				i	. !	MTPLTMGNYVFPKWGQGVGWLMALSSMVL IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV
]						RPENGPEQPQAGSSTSKEAYI
726	2076	Α	5711	156	423	PRRDPGRTPELRGSAPRKTGANMPVRRGHVA
						PONTFLGTIIRKFEGONKKFIIANARVONCAII
						YCNDGFCEMTGFSRPDVMQKPCTCD
727	2077	Α	5716	3	274	HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP
						LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL
728	2078	A	5737	1899	649	AWFEKMTCYLQLLFNICLPDVSEE
	2070	^ ]	3/3/	1633	049	IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEEISLGPACWLWDFLRRSQQAGFL
	٠	l		(		LPLSGGVDSAATACLIYSMCCQVCEAVRSGN
		l		ĺ		EEVLADVRTTVNQISYTPQDPRDLCGRILTTC
		ŀ			•	YMASKNSSQETCTRARELAQQIGSHHISLNID
	1				1	PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL
	İ	l	· ]	· [	l.	ALQNVQARIRMVLAYLFAQLSLWSRGVHGG
	ŀ	j		l	1	LLVLGSANVDESLLGYLTKYDCSSADINPIGG
			ŀ	ļ	. }	ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE
	l	ľ	l	- 1	1	LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD
	1	l		I	ļ	KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE
	ļ	J		I		DNRFDLRPFLYNTSWPWQFRCIENQVLQLER
	<u></u> l					AEPQSLDGVD
729	2079	A	5741	1	5976	PGCAARLSRARAPGPGAAGAGRKRLADPGPP
	ł				•	PASRRLRAPGSRPRLAPCTRRAAQPAHARMA
	ŀ					PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL
	}	l	Į	}	ł	LLLLLGAARAGALEIQRRFPSPTPTNNFALDG
. ]	1		1	. ]	l	AAGTVYLAAVNRLYQLSGANLSLEAEAAVG
. ]				ļ	l	PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV
ĺ					1	AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS
					ŀ	TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF
						PRNRSLEDHRFENTPEIAIRSLDTRGDLAKLFT
				L		I WHO THE THIRD TO I KOD LAKE I

HPGEPLTLVIHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIIHCSVNESLGAAVO LPITIQVGNFNQTIATLQLGGSETAIIVSIVICS LLLLSVVALFVFCTKSRRAERYWQKTLLQM EMESQIREEIRKGFAELQTDMTDLTKELNRS GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQ LNSQGSSQAQETHPLLGEWKIPESCRPNMER GISLFSSLLDNKHFLLVFVHALEQQKDFAVRI RCSLASLLTIALHGKLEYYTSIMKELLVDLII ASAAKNPKLMLRRTESVVEKMLTNWMSICI YSCLRETVGEPFFLLLCAIKQQINKGSIDAITO KARYTLNEEWLLRENIEAKPRNLNVSFQGC MDSLSVRAMDTDTLTQVKEKILEAFCKNVP SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKI NTLGRVKDLDTEKYFHLVLPTDELAEPKKSI RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDL KAILSIREDKPPLAVKYFFDFLEEQAEKRGISI PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDI TDHIDACLSVIAQAFIDACSISDLOLGKDSPT	SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  FDLNPSDDNILKIKQGAKEQHKLGFVSAFLHI SDPPPGAQSYAYLALNSEARAGDKESQARSL LARICLPHGAGGDAKKLTESYIQLGLQCAGG AGRGDLYSRLVSVFPARERLFAVFERPQGSP, ARAAPAALCAFRFADVRAAIRAARTACFVEP APDVVAVLDSVVQGTGPACERKLNIQLQPEG LDCGAAHLQHPLSILQPLKATPVFRAPGLTSV AVASVNNYTAVFLGTVNGRLLKINLNESMQ VVSRRVVTVAYGEPVHHVMQFDPADSGYLY LMTSHQMARVKVAACNVHSTCGDCVGAAD AYCGWCALETRCTLQQDCTNSSQQHFWTSA SEGPSRCPAMTVLPSEIDVRQEYPGMILQISGS LPSLSGMEMACDYGNNIRTVARVPGPAFGHC IAYCNLLPRDQFPFPPPNQDHVTVEMSVRVN GRNIVKANFTTYDCSRTAQVYPHTACTSCLSA QWPCFWCSQQHSCVSNQSRCEASPNPTSPQD CPRTLLSPLAPVPTGGSQNILVPLANTAFFQG AALECSFGLEEIFFAVWVNESVVRCDQVVLHTRKSQVFPLSLQLKGRPARFLDSPEFMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGG RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRTDTSI ACTMPEGALPAPVPCVCVFFERRGCVHGNLTF WYMQNPVITAISPRRSPVSGGRTTTVAGERFI MYQNVSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH
RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDL KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDI TDHIDACLSVIAQAFIDACSISDLOLGKDSPTI		·					LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRTFFFKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD
MNAHLAFESRKYQNEFNTNVAMAEIYKYAI RYRPQIMAALEANPTARRTQLQHKFEQVVA MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLOLLRTAOLPEOVSWPWGC	730	2080	A	5744	3		RQSHRKKVLPEIYLTRLLSTKGILQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTM KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEOVYAL
LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT   LKD     731   2081   A   5747   1   382   FLKCMRKAFRSSKLLQVGYTPDGKDDYRWG	731	2081	A	5747	1	382	LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LSFSLRSSRVSGRHWKNFALVPLLREASARD RQSAQPEEVYLRQFSGSLKPEDAEVFKSPAAS
732	2082	A	5753	198	3	GEK AQAESSTVASPEATAGPLCTRIPNVPPPTPIRP
733	2083		5754	2	2223	PGKLQAQLPCPSPVRFTSARIPPASRPQTKS  AAGPPGLEAEGRAPESAGPGPGGDAAETPGL PPAHSGTLMMAFRDVTVQIANQNISVSSSTAL SVANCLGAQTVQAPAEPAAGKAEQGETSGR EAPEAPAVGREDASAEDSCAEAGASGAADG ATAPKTEEEEEEETAEVGRGAEAEAGDLEQ LNRTSTSTKSAKSGSEASAASKDALQAMILS LPRYHCENPASCKSPTLSTDTLRKRLYRIGLN LFNINPDKGIQFLISRGFIPDTPIGVAHFLLQRK GLSRQMIGEFLGNSKKQFNRDVLDCVVDEM DFSSMELDEALRKFQAHIRVQGEAQKVERLIE AFSQRYCMCNPEVVQQFHNPDTIIFILAFAIILL NIDMYSPNIKPDRKMMLEDFIRNLRGVDDG ADIPRELVVGIYERIQQKELKSNEDHVTYVTK VEKSIVGMKTVLSVPHRRLVCCSRLFEVTDV NKLQKQAAHQREVFLFNDLLVILKLCPKKKS SSTYTFCKSVGLLGMQFQLFENEYYSHGTILV TPLSGSEKKQVLHFCALGSDEMQKFVEDLKE SIAEVTELEQIRIEWELEKQQGTKTLSFKPCGA QGDPQSKQGSPTAKREAALRERPAESTVEVSI HNRLQTSQHNSGLGAERGAPVPPPDLQPSPPR QQTPPLPPPPPTPPGTLVQCQQIVKVIVLDKPC LARMEPLLSQALSCYTSSSDSCGSTPLGGPG SPVKVTHQPPLPPPPPPYNHPHQFCPPGSLLH
734	2084	A	5788	8	362	GHRYSSGSRSLV SSVMGDLVGQGLEEQIVARDENSWLIDGGTP IDDVMRVLDIDEFPQSGNYETIGGFMMFMLR KIPKRTDSVKFAGYKFEVVDIDNYRIDQLLVT RIDSKATALSPKLPDAKDKEESVA
735	2085	A	5827		1257	MVFSAVLTAFHTGTSNTTFVVYENTYMNITL PPFFQHPDLSPLLRYSFETMAPTGLSSLTVNST AVPTTPAAFKSLNLPLQITLSAIMIFILFVSFLG NLVVCLMVYQKAAMRSAINILLASLAFADM LLAVLNMPFALVTILTTRWIFGKFFCRVSAMF FWLFVIEGVAILLISIDRFLIIVQRQDKLNPYR AKVLIAVSWATSFCVAFPLAVGNPDLQIPSRA PQCVFGYTTNPGYQAYVILISLISFFIPFLVILY SFMGILNTLRHNALRIHSYPEGICLSQASKLGL MGLQRPFQMSIDMGFKTRAFTTILILFAVFIVC WAPFTTYSLVATFSKHPYYQHNFFEISTWLL WLCYLKSALNPLIYYWRIKKFHDACLDMMP KSFKFLPQLPGHTKRRIRPSAVYVCGEHRTVV
736	2086	A	5870	3	268	FTRSDELARHYRTHTGEKRFSCPLCPKQFSRS DHLTKHARRHPTYHPDMIEYRGRRRTPRIDPP LTSEVESSASGSGPGPAPSFTTCL
737	2087	A	5871	2	521	LTWPQLFLETLPELLHMSRPAEDGPSPGALVR RSSSLGYISKAEEYFLLKSRSDLMFEKQSERH GLARRLTTARRPPASSEQAQQELFNELKPAV DGANFIVNHMRDQNNYNEEKDSWNRVART VDRLCLFVVTPVMVVGTAWIFLQGVYNQPPP QPFPGDPYSYNVQDKRFI
738	2088	A	5881	1	1160	LVVTAITAILAFPNEYTRMSTSELISELFNDCG LLDSSKLCDYENRFNTSKGGELPDRPAGVGV YSAMWQLALTLILKIVITIFTFGMKIPSGLFIPS MAVGAIAGRLLGVGMEQLAYYHQEWTVFNS

SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1.00	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
cotide	seq-	]	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
j	}	]	1	peptide		/=possible nucleotide deletion, \=possible
ļ		├	<del> </del>	sequence		nucleotide insertion
				j		WCSQGADCITPGLYAMVGAAACLGGVTRMT VSLVVIMFELTGGLEYIVPLMAAAMTSKWVA
1		ł	ł			DALGREGIYDAHIRLNGYPFLEAKEEFAHKTI
	1	[			· ·	AMDVMKPRRNDPLLTVLTQDSMTVEDVETII
1	l	1	[			SETTYSGFPVVVSRESORLVGFVLRRDLIISIE
						NARKKQDGVVSTSIIYFTEHSPPLPPYTPPTLK
		ſ	<b>i</b>	i		LRNILDLSPFTVTDLTPMEIVVDIFRKLGLRQC
1		1			•	LVTHNGRLLGIITKKDVLKHIAQMANQDPDSI
F00	2000	ļ	5000			LFN .
739	2089	A	5892	2	916	TLQLAASVPFFAISLISWWLPESARWLIINGKP
ł		Ì	-			DQALQELRKVARINGHKEAKNLTIEVLMSSV
1						KEEVASAKEPRSVLDLFCVPVLRWRSCAMLV     VNFSLLISYYGLVFDLQSLGRDIFLLOALFGA
l i			1	-		VDFLGRATTALLLSFLGRRTIQAGSQAMAGL
						AILANMLVPQDLQTLRVVFAVLGKGCFGISL
						TCLTTYKAELFPTPVRMTADGILHTVGRLGA
1						MMGPLILMSRQALPLLPPLLYGVISIASSLVVL
			ŀ			FFLPETQGLPLPDTIQDLESQKSTAAQGNRQE
	2000					AFTVESTSLLEIVALHGAL
740	2090	A	5900	2	426	RPIKTLGIGFHFSVDGVHFLTQREVQNLWKE
				i		NLIILDTAKKHGYEVVDTFTTTMGRYKEFLQG KCGCHFHEVVKSKLSKEYNFIKMKRSRNHIM
]				J		GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV
						CSEILLSRMCANKRTM
741	2091	Α	5910	3	412	RMPESTLLIICENGYILEAPLPTIKOEEDDHDV
						VSYEIKDMCIKCFHFSSVKSKILRLIEIEKRER
-						QRELKEKIREERRNKLAAEMGEDGEKEFQEE
		1	ĺ	j		EEEKEEEEEEPLPEIFIPSTPSPILCGFYSEPG
L	0000	, -	5025			KFWV
742	2092	A	5936	1	482	MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV
		ĺ	[	I		TQMGNDKSIKCEQNLGHDTMYWYKQDSKK FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL
'		1	1	ŀ		NLHINSLELGDSAVYFCASSQDTALQSHCIPV
<b>!</b>		. 1				HKPPGSARKLQGSVCTCTQGSSLHSLMASDG
	1	ľ	- 1	ĺ		VPVC
743	2093	A	5938	1	1566	MNSFFGTPAASWCLLESDVSSAPDKEAGRER
	[					RALSVQQRGGPAWSGSLEWSRQSAGDRRRL
		J	1	l		GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA
		j	}	Į.	Į	ADRARRERFIMNEKWDTNSSENWHPIWNVN
		- 1	l.	j	j	DTKHILYSDINITYVNYYLHQPQVAAIFIISYF
		}		ŀ	[	LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI
	1	ļ				LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRFQCVVY
	Ì	Ī			. 1	PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH
	· 1					VQEEKYYRVRLNSQNKTSPVYWCREDWPNO
		İ	j	1		EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF
			j			RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI
	[	- 1	į	1	[	VALLFILSWLPLWTLMMLSDYADLSPNELQII
		- 1		ľ		NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG
	l	- 1		I	1	FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN
	İ	- 1		- 1		TSNQLVQESTFQNPHGETLLYRKSAEKPQQE
744	2004		5066	140	207	LVMEELKETTNSSEI
744	2094	A	5966	149	327	SHVCVSHYAGSSGCPAGAGAGAVALGISAVA
745	2095	${\mathbf{A}}$	5970	413	856	LYDYQGGRLGVARGAWYMEAPDIRQGDM
175	2073	^	2210	713	0.00	GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV
		1	- [	l	ļ	CGALMWIMLILVGLGFPFIMEALSHFLYVPFL
		-	l	į		GVCVCGAIYTGLFLPEIKGKTFQEISKELHRL
		!		}	ļ	NFPRRAQGPTWRSLEVIQSTEL

CEO TO	I CEO TO	N/~	1 000	Dendine d	Dung 27 - 4 - 3	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	neuce	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1.	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	i	1	ĺ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	l .	1		peptide		/=possible nucleotide deletion, \=possible
İ	1	1	i	sequence	l	nucleotide insertion
746	2096	A	5971	3	1343	AQTARRIIGLELDTEGHRLFVAFSGCIVYLPLS
1			ł			RCARHGACQRSCLASQDPYCGWHSSRGCVD1
	ł	1	ı		İ	RGSGGTDVDQAGNQESMEHGDCQDGATGSQ
		1	į			SGPGDSAYGVRRDLPPASASRSVPIPLLLASV
1	j	1				AAAFALGASVSGLLVSCACRRAHRRRGKDIE
1			1			TPGLPRPLSLRSLARLHGGGPEPPPPSKDGDA
1	1	İ	1	ł		VQTPQLYTTFLPPPEGVPPPELACLPTPESTPE
1		1	-			
		ı	1		ł	LPVKHLRAAGDPWEWNQNRNNAKEGPGRSR
	i	1	1		İ	GGHAAGGPAPRVLVRPPPPGCPGQAVEVTTL
	l	1	ŀ			EELLRYLHGPQPPRKGAEPPAPLTSRALPPEP
1	1	1	ł	1	ł	APALLGGPSPRPHECASPLRLDVPPEGRCASA
	[	]	1			PARPALSAPAPRLGVGGGRRLPFSGHRAPPAL
1	1 .	1				LTRVPSGGPSRYSGGPGKHLLYLGRPEGYRG
1	]	l	1			RALKRVDVEKPQLSLKPPLVGPSSRQAVPNG
-		<u> </u>	L.			GRFNF
747	2097	Α	5998	2	754	DHASLPCSWNHRFDVETRHVFIGDHSGQVTI
1		1				LKLEQENCTLVTTFRGHTGGVTALCWDPVQ
1		1	1 .			RVLFSGSSDHSVIMWDIGGRKGTAIELQGHN
l .						DRVQALSYAQHTRQLISCGGDGGIVVWNMD
f	ſ	ĺ	[		_	VERQETPEWLDSDSCQKCDQPFFWNFKQMW
		l	Ì :			DSKKIGLRQHHCRKCGKAVCGKCSSKRSSIPL
	1					MGFEFEVRVCDSCHEAITDEERAPTATFHDSK
1 .	ŀ	1	1 1	i		HNIVHVHFDATRGWLLTSGTDKVIKLWDMT
	l		<u> </u>			PVVS
748	2098	Α	6001	2	747	AMVFGGVVPYVPQYRDIRRTQNADGFSTYV
ļ.			1			CLVLLVANILRILFWFGRRFESPLLWQSAIMIL
						TMLLMLKLCTEVRVANELNARRRSFTAADS
l	1	i	1 1	i	*:	KDEEVKVAPRRSFLDFDPHHFWQWSSFSDYV
	}	l				QCVLAFTGVAGYITYLSIDSALFVETLGFLAV
1	ł ·	1	1 1			LTEAMLGVPQLYRNHRHQSTEGMSIKMVLM
)			}			WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV
1			1 1			DLAILGQAYAFARHPQKPAPHAVHPTGTKAL
749	2099	Α	6002	2	447	GRPDRSELVRMHILEETFAEPSLQATQMKLK
				_		RARLADDLNEKIAORPGPMELVEKNILPVDSS
]				ĺ		VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP
		l ,		J		DOPASQESOGSAASPSEPKVSESPSPVTTNTP
				ŀ		AQFASVSPTVPEFLKTPPTAD
750	2100	Α	6004	2	427	LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG
'	2.50		5007	-	741	
				ł		WRWELRLRNYVPEDEDLNKRRVPQAKPDAV QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD
	·				×	
						WDLKRDVAKKLEKLLKRTQRAIAELIRERLK
751	2101	A	6000	33	1200	GQEDSLDSAVDAATEHKTC
751	2101	A	6007	33	1280	TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF
				İ	·	SHPDKLKRMSKSVPAFLQDESDDRETDTASE
		į			ļ	SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS
			ŀ	1		VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL
	j	j	i j	j	J	HVFVAQCKDLAAADVKKQRSDPYVKAYLLP
! <b>!</b>				ŀ		DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK
	·			}		QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE
' <b>i</b>	ļ		İ	1		TWDWDNKQNKQLRWYPLKRKTAPVALEAE
' I	l	- 1	1	•		NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV
	ľ	Į			j	KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ
					į	KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC
ĺ	i	ı	1	i		VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT
	ľ	ŀ	j	į.		EVDWMDSTSEEVALWEKMVNSPNTWIEATL
	ł	ł	1	ļ	1	PLRMLLIAKISK
752	2102	A	6028	108	1283	KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL
İ						SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF
j		i			į	AAAIPGHRCWVHMLDNNTGSGNETGILSEDA
						THE STREET THE STREET STREET STREET STREET

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutarnic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ŀ		.		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
	-					LLRISIPLDSNLRPEKCRRFVHPQWQLLHLNG
1			1			TIHSTSEADTEPCVDGWVYDQSYFPSTIVTKW
1		ľ				DLVCDYQSLKSVVQFLLLTGMLVGGIIGGHV
						SDRFGRRFILRWGLLQLAITDTCAAFAPTFPV
						YCVLRFLAGFSSMIIISNNSLPITEWIRPNSKAL
1						VVILSSGALNIGQIILGGLAYVFRDWQTLHVV
<b>l</b>		'				ASVPFFVFFLLSRWLVESARWLIITNKLDEGL
						KALRKVARTNGIKNAEETLNIEVVRSTMQEE
1						LDAAQTKTTVWDLFRNPSMRKRICILVFLRK
753	0100		60.43		1450	KNLKEKA
753	2103	A	6043	1	1470	DSFESILRLIFEIHHSGEKGDIVVFLACEQDIEK
						VCETVYQGSNLNPDLGELVVVPLYPKEKCSL
		ĺ		[		FKPLDETEKRCQVYQRRVVLTTSSGEFLIWSN
	•					SVRFVIDVGVERRKVYNPRIRANSLVMQPISQ
						SQAEIRKQILGSSSSGKFFCLYTEEFASKDMTP LKPAEMQEANLTSMVLFMKRIDIAGLGHCDF
						MNRPAPESLMQALEDLDYLAALDNDGNLSE
1						FGIIMSEFPLDPQLSKSILASCEFDCVDEVLTIA
						AMVTAPNCFSHVPHGAEEAALTCWKTFLHPE
1						GDHFTLISIYKAYODTTLNSSSEYCVEKWCRD
						YFLNCSALRMADVIRAELLEIIKRIELPYAEPA
				•		FGSKENTLNIKKALLSGYFMOIARDVDGSGN
1						YLMLTHKQVAQLHPLSGYSITKKMPEWVLF
1						HKFSISENNYIRITSEISPELFMQLVPQYYFSNL
						PPSESKDILQQVVDHLSPVSTMNKEQQMCET
						CPETEQRCTLQ
754	2104	Α	6055	2	394	YYALHHWPFPDLLCQTTGAIFQMNMYGSCIF
						LMLINVDRYAAIVHPLRLRHLRRPRVARLLC
						LGVWALILVFAVPAARVHRPSRCRYRDLEVR
1						LCFESFSDELWKGRLLPLVLLAEALGFLLPLA
						AVVYSS
755	2105	A	6059	3	1795	LGLGSGTLLSVSEYKKKYREHVLQLHARVKE
í I		1				RNARSVKITKRPTKLLIAPESAAPEEALGPAEE
]					ì	PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVL
i l		1				QGPAGIGKTMAAKKILYDWAAGKLYQGQVD
						FAFFMPCGELLERPGTRSLADLILDQCPDRGA
						PVPQMLAQPQRLLFILDGADELPALGGPEAAP
				•		CTDPFEAASGARVLGGLISKALLPTALLLVTT
						RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK
				ŀ	<u></u>	YFYKFFRDERRAERAYRFVKENETLFALCFV
				; <u> </u>		PFVCWIVCTVLRQQLELGRDLSRTSKTTTSVY
			·	٠. ا		LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVOTLFLSK
						KELPGVLETEVTYQFIDQSFQEFLAALSYLLE
					·	DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT
						RFLFGLLSAERMRDIERHFGCMVSERVKQEA
		' I	1	ł	l	LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE
				ļ		PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA
'				ļ		LCRFPELALQRVRFCRMDVAVLSYCVRCCPA
				1		GQALRLISCRLVAAQEKKKKSLGKRLQASLG
				ļ		GG
756	2106	A	6060	12	436	SGRPTRPAKPTGQGMGRFMLTLVCQGSIMMS
""	2100	A	5000		430	ARDLIMNNLTELOPGLIPHILIFLEELRLSGNH
				J	ļ	LSHIPGOAFSGLYSLKILMLHNNOLGGIPAOA
ì			•			LWELPSLQSLRLDANLISLVPERSFEGLSSLRH
		.			. (	LWLDDNALTEIPS
757	2107	A	6063	54	419	ITPLGLGAADMCAFPWLLLLLLQEGSQRRL
		**		-	-17	WRWCGSEEVVAVLQESISLPLEIPPDEEVENII
						WSSHKSLATVVPGKEGHPATIMVTNPHYQG
		l				

SEQ ID	SEQ ID	Met	SEQ	Predicted	Desdicted on 3	I Amino ocid company (A. Ali i G. G.
NO: of	NO: of	hod	ID NO:	beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid,
cotide	seq-	ł	USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	]	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	i	l		peptide		/=possible nucleotide deletion, \=possible
1				sequence	Ì	nucleotide insertion
						QILTMLLRSLQQPSASWPRDCSSSCSW
758	2108	A	6066	125	438	IGISCPATIFVPMFSHSLIGIGEEYQLPYYNMV
1	j	l		ŀ		PSDPSYEDMREVVCVKRLRPIVSNRWNSDEC
1		1				LRAVLKLMSECWAHNPASRLTALRIKKTLAK
						MVESQDVKI
759	2109	A	6072	3	650	PGRRFRPAALEERAMEKLREKVPFQNRGKGT
		ĺ				LSSIIPNNSDTRKATETTSLSSKPEYVNPDFRW
	1					SKDPSSKSGNLLETSEVGWTSNPEELDPIRLA
l i	İ					LLGKSGLSCQVGSATSHPVSCQEPIDEDQRISP
		1				KDKSTAGREFSGQVSHQTTSENQCTPIPSSTV
}	1	l				HSSVADMQNMPAAVHALLTQPSLSAAPFAQ
760	2110	<u> </u>				RYLGTLPSTGSTTLPQCHAGNATVW
<sup>760</sup>	2110	A	6077	3	730	PLRLTLMEEVLLLGLKDREGYTSFWNDCISSG
}						LRGCMLIELPLRGRLQLEACGMRRKSLLTRK
1			ĺ			VICKSDAPTGDVLLDEALKHVKETQPPETVQ
1.						NWIELLSGETWNPLKLHYQLRNVRERLAKNL
1						VEKGVLTTEKQNFLLFDMTTHPLTNNNIKQR
1 .						LIKKVQEAVLDKWVNDPHRMDRRLLALIYL
1 1						AHASDVLENAFAPLLDEQYDLATKRVRQLLD LDPEVECLKANTNEVLWAVVAAFTK
761	2111	A	6078	833	390	IVSFHLSGFKKFVRPFSFLSVHGLQVDEYHSV
			0070	~ l	370	HQKLSADMADHSNLIRSLLVGAEDARLMRD
1				j		MKTMKSRYMELYDLNRDLLNGYKIRWNNH
						TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT
						ACRDAIRSNNINTLFKIMRVGTASS
762	2112	Α	6079	2	2686	KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS
1						HSDLWSSSSSLESSSFPLPKQYLDVSSQTDISG
	ł					SFGINSNNQLAEKVRLRLRYEEAKRRIANLKI
i 1						QLAKLDSEAWPGVLDSERDRLILINEKEELLK
Į [	[					EMRFISPRKWTQGEVEQLEMARKRLEKDLQ
1 1	i i					AARDTQSKALTERLKLNSKRNQLVRELEEAT
	ı	-				RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR
<b>!</b> !	ŀ	i				GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ
]					j	SKVEFLLLEGATGFRPSGCITTIHEDEVAKTQ
}			ĺ			KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS
	ļ					PSPPCSPLMADPLLAGDAFLNSLEFEDPBLSA
T-	1	- 1			ļ	TLCELSLGNSAQERYRLEEPGTEGKQLGQAV
ĺ	ļ					NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIOIALK
	İ	i	ľ	ſ	- 1	YDEKNKQFAILIIQLSNLSALLQQQDQKVNIR
				1		VAVLPCSESTTCLFRTRPLDASDILVFNEVFW
	ļ	l	ļ	1		VSMSYPALHQKTLRVDVCTTDRSHLEECLGG
	ŀ	ŀ	1	. }		AQISLAEVCRSGERSTRWYNLLSYKYLKKQS
				l		RELKPVGVMAPASGPASTDAVSALLEQTAVE
		l		l	İ	LEKRQEGRSSTQTLEDSWRYEETSENEAVAE
	1	·		l	J	EEEEEVEEEGEEDVFTEKASPDMDGYPALK
	1	J		J	1	VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF
		j	- 1	ľ	ł	LRGSTURSKTFSPGPQSQYVCRLNRSDSDSST
		1	1	l	l	LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK
				į	l	SLRSERLIRTSLDLELDLQATRTWHSOLTOEIS
1	-	l	- 1	į	` '	VLKELKEQLEQAKSHGEKELPQWLREDERFR
1		1	1	ļ	ŀ	LLLRMLEKRMDRAEHMGELQTDKMMRAAA
1		I		1	. [	KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR
7(2)		,				MNIPALSADDV
763	2113	A	6082	3	1558	PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE
	1				.	VLENLTQGKMCLVPGKTRKLLFKFVAKTED
1						VGKKIEITSVDLALGNETGRCVVLNWQGGGG
				į		DAASSQEALQAARSFKRRPKLPDNEVHWGSII
						IQASTMIISRVPNISVHLLHEPPALTNEMYCLV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  VTVQSHEKTQIRDVKLTAGLKPGQDANLTQK THVTLHGTELCDESYPALLTDIPVGDLHPGEQ
	·			-		LEKMLYVRCGTVGSRMFLVYVSYLINTTVEE KEIVCKCHKDETVTIETVFPFDVAVKFVSTKF EHLERVYADIPFLLMTDLLSASPWALTIVSSE LHLAPSMTTVDQLESQVDNVILQTGESASECF CLQCPSLGNIEGGVATGHYIISWKRTSAMENI PIITTVITLPHVIVENIPLHVNADLPSFGRVRES LPVKYHLQNKTDLVQDVEISVEPSDAFMFSG LKQIRLRILPGTEQEMLYNFYPLMAGYQQLPS LNINLLRFPNFTNQLLRRFIPTSIFVKPQGRLM DDTSIAAA
764	2114	A	6093		1422	AAADIANSNAGAAVGRKAGPRSPPSAPAPAP PPPAPAPPTIGNNHQESPGWRCCRPTLRERN ALMENNELMADVHFVVGPPGATRTVPAHKY VLAVGSSVFYAMFYGDLAEVKSEIHIPDVEPA AFLILLKYMYSDEIDLEADTVLATLYAAKKYI VPALAKACVNFLETSLEAKNACVLLSQSRLF EEPBLTQRCWEVIDAQAEMALRSEGFCEIDR QTLEIIVTREALNTKEAVVFEAVLNWAEAEC KRQGLPITPRNKRHVLGRALYLVRIPTMTLEE FANGAAQSDILTLEETHSIFLWYTATNKPRLD FPLTKRKGLAPQRCHRFQSSAYRSNQWRYRG RCDSIQFAVDRRVFIAGLGLYGSSSGKAEYSV KIELKRLGVVLAQNLTKFMSDGSSNTFPVWF EHPVQVEQDTFYTASAVLDGSELSYFGQEGM TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE LIFYA
765	2115	A		1	1150	SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF RPVKAPGTFHMVHGKCMCKHNTAGSHCQH CAPLYNDRPWEAADGKTGAPNECRTCKCNG HADTCHFDVNVWEASGNRSGGVCDDCQHN TEGQYCQRCKPGFYRDLRRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK
766	2116	A	6103	2	384	MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS
	2117	A	6106	1	542	SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW
768	2118	A	6109	3	292	FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE
769	2119	A	6110	1	711	RHEPSCSNGVASTKSKONHSKYPAPSSSSSS SSSSSSSSSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY KHEDLQTDESSMDDRHPRRQLCGGNQAATE

SEQ ID	SEQ ID	Met	Lego	/ D3:4-3	1 D 3 - 1 - 3	
NO: of	NO: of	•	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
		hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F-Phenylalanine, G-Glycine, H-Histidine,
cotide	seq-	J	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	ł	914	ng to first	acid residue	Q=Ghitamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	J			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			peptide	30quante	/=possible nucleotide deletion, \=possible
	ł	l				/-possible nucleonde defetion, /-possible
	<del></del>	<del> </del>	<del> </del>	sequence		nucleotide insertion
	1		l			RIILFGRELQALSEQLGREYGKNLAHTEMLQD
1	ł .	l	ļ	ł	ł	AFSLLAYSDPWSCPVGQQLDPIQREPVCAAL
	ſ	ļ	1	1		NSAILESQNLPKQPPLMLALGQASECLRLMA
	i		i	1		RAGLGSCSFARVDDYLH
770	2120	A	6125	2	570	YFGLNLHVQHLGNNVFLLQTLFGAVILLANC
ł	l			} <del>-</del>	1	VAPWALKYMNRRASQMILMFILAICLIAIIF
1	ľ		1			VALVALA IMIVAÇAS QWILLMIT LAICLLAIIF
1		1	1	i		VPQEMQMLREVLATLGLGASALANTLAFAH
ı	i	1	ŀ		1	GNEVIPTIIRARAMGINATFANIAGALAPLMM
1	1		ł .		ŀ	ILSVYSPPLPWIIYGVFPFISGFAFLLLPETRNK
	<u> </u>	L				PLFDTIQDEKNERKDPREPKQEDPRVEVTQF
771	2121	Α	6126	909	353	RSFVLDTASAICNYNAHYKNHPKYWCRGYF
1		Į	]			RDYCNIIAFSPNSTNHVALRDTGNQLIVTMSC
1	i	1	] .			LTKEDTGWYWCGIQRDFARDDMDFTELIVT
1						DDKGTLANDFWSGKDLSGNKTRSCKAPKVV
		1				DEVENTANT MOUNT CHOINE MOUNT OF THE MOUNT OF
1		ĺ	i i	1		RKADRSRTSILIICILITGLGIISVISHLTKRRRS
772	2122	A	6148		-0.0	QRNRRVGNTLKPFSRVLTPKEMAPTEQM
172	2122	A	0148	7	810	FVLGILALSHTISPFMNKFFPASFPNRQYQLLF
						TQGSGENKEEIINYEFDTKDLVCLGLSSIVGV
1 1	İ	İ	!			WYLLRKHWIANNLFGLAFSLNGVELLHILNN
} !						VSTGCILLGGLFIYDVFWVFGTNVMVTVAKS
	•		j l			FEAPIKLVFPQDLLEKGLEANNFAMLGLGDV
			1			VIPGIFIALLLRFDISLKKNTHTYFYTSFAAYIF
1 1						GLGLTIFIMHIFKHAQPALLYLVPACIGFPVLV
1		i				ALAKGEVTEMFSYEESNPKDPAAVTESKEGT
773	2123	A	6161	3	1088	EASASKGLEKKEK
1 '''	لكلك	n.	0101	3	1088	CQPMLVTRKNHPKLLLRRTESVAEKMLTNW
1						PTFLLYKFLKESAGEPLPMLYCAIKHQMEKG
1 1						PIDATTGEARYSLSEDKLIRHLIDYKTLTLNCV
1 1						NPENENAPEVPVKGLDCDTGTQAKEKLLDA
j j						AYKGVPYSQRPKAADMDLEWRQGRMARIIL
1 1					,	QDEDVTTKIDNDWKRLNTLAHYQVTDGSSV
1						ALVPKQTSAYNISNSSTFTKSLSRYESMLRTA
1 1				•		SSPDSLRSRTPMITPDLESGTKLWHLVKNHDH
) )			i j			LDQREGDRGSKMVSEIYLTRLLATKGTLQKF
	1		ļ .	i	· ·	LDQREODROSKMIVSEITLIRGLAIRGILQKF
	İ	اء		ļ	,	VDDLFETIFSTAHRGSALPLAIKYMFDFLDEQ
1				İ		ADKHQIHDADVRHTWKSNCLPLRFWVNVIK
1224	2124				·	NPQFVFDIHKNSITDACLSVV
774	2124	A	6163	860	125	KTAVKKRNLNPVFNETLRYSVPQAELQGRVL
1 1					. [	SLSVWHRESLGRNIFLGEVEVPLDTWDWGSE
1 1					•	PTWLPLQPRVPPSPDDLPSRGLLALSLKYVPA
j i				1		GSEGAGLPPSGELHFWVKEARDLLPLRAGSL
; [	ſ	. 1		ľ	ľ	DTYVQCFVLPDDSRASRQRTRVVRRSLSPVF
	l	1		.	•	NHTMVYDGFGPADLRQACAELSLWDHGALA
	- 1			. I		
1 I	1					NRQLGGTRLSLGTGSSYGLQVPWMDSTPEEK
775	2125	<del>,  </del>	(10)			QLWQALLEQPCEWVDGLLPLRTNLAPRT
'' <sup>2</sup> .	2125	A	6191	2	392	ARGIGSLGRDHSGSGGGTGMAGAWVRKAAD
j i		- 1	.	. [		YVRSKDFRDYLMSTHFWGPVANWGLPIAAIT
j l	j	]		I		DMK\KSPEIISRRMTFAL*CYSLTFVRFAHYVQ
1 1	Į.	j	j	J		PWNWLMLGCHTAVDFDQLISSMPCISHGMT
, · I						ASASAL CISHOWI
776	2126	A	6217	1	827	FRGYWGVREAFTDASWSGGLGPGKPGMKIT
			~·	- !	·-·	DOVI A VEHI CHEND THE TOTAL THE TOTA
1 1	Į.	ļ	- 1	Ì	J	RQKHAKKHLGFFRNNFGVREPYQILLDGTFC
	ŀ	1		i		QAALRGRIQLREQLPRYLMGETQLCTTRCVL
]	l	J	}	l		KELETLGKDLYGAKLIAQKCQVRNCPHFKNA
[		j		l		VSGSECLLSMVEEGNPHHYFVATQDQNLSVK
į į	j	J	J	l	· . [	VKKKPGVPLMFIIQNTMVLDKPSPKTIAFVKA
}						VESGIRLSQCMRKKVSNISKRNRV**KTLNRG
	ŀ		1		į	RRKKRKKISGPNPLSCLKKKKKAPDTQSSASE
ı İ	1	- 1	İ	i		KKRKRKRIRNRSNPKVLSEKQNAEGE
						THE PROPERTY AND THE PROPERTY OF THE PROPERTY

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino goid cognomes (4-41-4-6-0-0-1
NO: of	NO: of	hod	IDNO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide	1 1100	in in	nucleotide		
eotide			1 —	1	location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
tience		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \-possible
				sequence	1	nucleotide insertion
777	2127	Α	6236	1038	1402	YYQISSLPSIVGNGIFLWLLICIFLAKQGGSRL*
i	<b>5</b> 1		i i		Ī	FQPFGRPRGGGHLRSGVLGQPGQHGETP/SFF
1			i			YNSKISPALWGPPVIPSALGGEAGKSL*PRRQ
	i I		1			RFQRGGIAPLPSRVRGRAKLFLKKK
778	2128	A	6237	422	913	ASFFHHHRGAFLLLLAIPGS*GQDQSLIHWSN
					7.20	AVSNAD\LLDLK\N*LDH\LEEKMPL\EVKVVP
1						PQVL\SEPN*RSGGCFSAPSFEVPPWTGEVKP/
						SPQRDGGALG\QGPLGIPSDSILALLKKOT*RA
i						LLNWPLGSLRRSSCFGGQDGQDLKPRSGLGC
	ĺ					NSFRYRR
779	2129	A	6249	420	26	
113	2129	A	0249	420	36	ARAPSPSFSVRDVELSDPARERGEMPVAVGP
						YGQSQPSCFDRVKMGFVMGCAVGMAAGAL
	i					FGTFSCLSSILVSSSG/SGMRGRELMGGIGKTM
						MQSGGTFGTFMAIGMGIRC*PWLPTTSVPSH
L						QSQPMY
780	2130	A	6263	415	1380	RIMRMCDRGIQMLITTVGAFAAFSLMTIAVG
						TDYWLYSRGVCRTKSTSDNETSRKNEEVMT
						HSGLWRTCCLEGAFRGVCKKIDHFPEDADYE
						QDTAEYLLRAVRASSVFPILSVTLLFFGGLCV
1						AASEFHRSRHNVILSAGIFFVSAGLSNIIGIIVYI
1						S\ANAGRTPGQR\DSKKSYSYGWSF/YFSGAFS
1						FIIGR/IIC*GVGLPWHIYIEKHQQLRAKSHSEF
			İ			LKKSTFARLPPYRYRFRRRSSSRSTEPRSRDLS
						PISKGFHTIPSTDISMFTLSRDPSKITMGTLLNS
I .						DRDHAFLQFHNSTPKEFKESLHNNPANRRTT
1 1						PV
781	2131	A	6274	832	318	RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH
] '''		*	02/1		316	LQVKQKQLACLCTWQARDPDCPPSTKVVL/L
1						COCHCCATALEODOLARIONICO MOOT DATO
1 .1						VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS
1 1						QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD
1 1						LEFLGDLKGCSELKNFQELITQSALVHPKADV
782	2122		×			WWYCGRPLLGTLPSN
102	2132	A	6281	1324	393	WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG
1 .						EREDWGIGSA*SVGAVSKVPSARF*RTYPS\E
1 1						DEEEVTHQKSSSSDSNSEEHRKKKTSRSRNK
1		- 1			ļ	KKRKNKSSKRKHRKYSDSDSNSESDTNSDSD
		1			İ	DDKKRVKAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK
	1					ESSDSSCKDSEEDLSEATWMEQPNVADTMDL
1 1					J	1GPEAPIIHTSQDEKPLKYGHALLPGEGAAMA
		- 1		l	ļ	EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM
		- 1			. [	SGSRHRRMEAVRLRKENQIYSADEKRALASF
						NQEERRKRESKILASFREMVHKKTKGKDDK
783	2133	A	6305	201	1032	WDDYPQGALRREAAEGLHFLGPPGRVRGQ
		(		j		LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP
1 1		ł				AMARPPVPGSVVVPNWHES/RRGQGVPGLHS
1 1		I				AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV
		ì		j		SGKSGVGKTALVAKLAGLEVPVVHHETTGIO
1 1					1	
j l				1		TTVVFWPAKLQASSRVVMFRFEFWDCGESA
				1		LKKFDHMLLACMENTDAFLFLFSFTDRASFE
		i				DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT
704	2124					DVPERDLTAFRQAWELPLLRVKSVPGRRLG
784	2134	A	6308	86	96	GSSPDPASLITMKNQDKKNGAAKQSNPKSSP
I	ł	1	1	ł		GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA
j i		1				PRKPEGAQARTAQSGALRDVSEELSRQLEDIL
] [		1			.	STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR
1		ļ	1			TYVARNGEPEPIPVVNGEKEPSKGDPNTEEIR
					:	QSDEVGDRDHRRPQEKKKAKGLGKEITLLM
						QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ
1	- 1	ŀ	}		}	KQMKLLQKKQSQLVQEKDHLRGEHSKAVLA
					1	ATA WYCHATCHOOL A CONTRACTOR OF THE CONTRACTOR O

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteme,
nucl-	peptide	100	in in	nucleotide		D=Aspartic Acid, E=Glutamic Acid,
cotide					location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	}		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ı	i	l ·		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1	1	peptide		possible nucleotide deletion, possible
1	ļ	į	1	sequence		nucleotide insertion
<del></del>	<del> </del>			Scharce	ļ	nucleotide insertion
1	į.		ļ	Į.		RSKLESLCRELQRHNRSLKEEGVQRAREEEE
ì	1	ļ	i	1		KRKEVTSHFQVTLNDIQLQMEQHNERNSKLR
Į.	1	)	ļ	]	i	QENMELAERLKKLIEQYELREEHIDKVFKHK
		l			i	DLQQQLVDAKLQQAQEMLKEAEERHQREKD
1						FLLKEAVESQRMCELMKQQETHLKQQLALY
1	1		1	•	1	TEKFEEFONTLSKSSEVFTTFKQEMEKMTKKI
1	1	t	ł		,	KKLEKEITMYRSRWESSNKALLEMAEEKTV
ļ	-					ARLENET IM I KOKWESSNKALLEMAEEK IV
	1	1				RDKELEGLQVKIQRLEKLCRALQT/GAQ*PVR
505						GQRWGSHRTSAVRIFS
785	2135	Α	6319	1493	889	SPQGPLLRSVSPVSAGASSVTPGGAQPGVTTT
1	}		,			PPSLVAVAPAPGSAAGPAAGWQ*HAGCR/WT
	1					KLPWSWGMRPMKIFFSEEYRSISTRISHDAL*
1						EKCTQPAKPLSMIR\TGSSVSPG/PLVKWNWT
į .	ļ ·	l				DDEEDFIGGERALICATION CANONICAL CANON
1		}				RREFRNSGTRVVSSCCGMSCMYSFLGHCSV/S
		·		į		QDLPLVHVDVGWQPPLGPTVGLRPGLLPLHD
506	200					TTPCQKLVVDDLDWA
786	2136	Α	6320	551	135	RWLPVAECDSSCVGCTGEGPGNCKECISGYA
						REHGQCADVDECSLAEKTCVRKNENCYNTP
1		1				GSYVCVCPDGFEET/RRCLCAAGRG*SHRRRK
1 1						PDTAALPRRPVMCRTYPLNYSEGCPVENVAL
1						RMPSPAVDSGGERLPAL
787	2137	A	6330	1693	227	
'''	2137	Λ	0530	1073	221	DYVLTAELHRQRSPGVSFGLSVFNLMNAIMG
[ ]	•					SGILGLAYVMANTGVFGFSFLLLTVALLASYS
1 1						VHLLLSMCIQTAYLGP*TNYFMVLPAH*LTCL
1						PLIEFLQSL*NSL\*AVTSYEDLGLFAFGLPGKL
			ı			VVAGTIIIQNIGAMSSYLLIIKTELPAAIAEFLT
i l						GDYSRYWYLDGQTLLIIICVGIVFPLALLPKIG
i i						FLGYTSSLSFFFMMFFALVVIIKKWSIPCPLTL
						NYVEKGFQISNVTDDCKPKLFHFSKESAYALP
] ]						TMAFSFLCHTSILPTYCELQSPSKKRMQNVTN
1 1					ì	TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE
	1		Į.		Î	VTCHRIKDKVESELLKG***IP*SHDVVVMT\V
	1		1			KLCILFAVLL\TVPLIHFPARKAVTMMFFSNFP
. I						FSWIRHFLITLALNIIIVLLAIYVPDIRNVFGVV
1 1	ſ		ſ			GASTSTCLIFIFPGLFYLKLSREDFLSWKKLGV
Ł I				1		GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI
788	2138	A	6351	1	6622	PRSLCFSLWAEAAVLADGGLRRRRRLLRGTM
				_		SASFVPNGASLEDCHCNLFCLADLTGIKWKK
1 1			1			VANOCINICA DE ENTREPENTE COMO ET EL ENT
1 1	1	ł	i	- 1	}	YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV
	}		ŀ			LG/VWRRDQRPERRE\L*IFWGGEDP\VLLTLF
l I		ł	l			TMTYQKKKMECGRMDFPMNAVLCFSKAVH
}	· 1		ľ	1		NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN
[	ľ	i	ì	. 1	I	KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY
		į	l	l		LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ
		i	Į	l		AFKMSDSATKKLIGEWKQFYPISCCLKEMSE
	I	ĺ	1	J	}	EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC
i l	I			l	1	EM ADOCDIMEDON/COMPAGENT AND AND A COMPAGENT A
, 1	ļ	J	)	J	1	FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS
			ľ		1	TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW
	}	j				VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV
	l	l				DRVWQECNMNRAQNKRKYSASSGGLCEEAT
! !	- 1	1	}			AAKVASWDFVEATORTNCSCLRHKNLKSRN
}	l			J	}	AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK
·	l	J	i	1	j	PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL
	l	- 1	. 1	i	1	MICADIDECTION TO TO TO TOTAL AND
	J	- 1	ł		l	VISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE
j j	}	l		l	l	MANSPQPPPLSPHPCDVVDEGVTKTPSTPQS
	í	- }	- 1	f	1	QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ
		i		ļ	l	YQEAVEPTVYVGTAVNLEEDEANIAWKYYK
ļ l		l	ļ	Ì	l	FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV
		I	}	J	į	TSVTELMVQCKKPLKVSDELVQQYQIKNQCL
		ſ		ſ	}	SAIASDAEQEPKIDPYAFVEGDEEFLFPDKKD

CONO	Compa has		1 6-6	-		
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of nucl-	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	seq- uence		USSN 09/496	location	to last amino	I=Isoleucine, K=Lysine, L=Leucine,
nence	ucucc	l	914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
		·		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
				peptide	sequence	possible nucleotide deletion. \=possible
		Į		sequence		nucleotide insertion
			T			RONSEREAGKKHKVEDGTSSVTVLSHEEDA
						MSLFSPSIKQDAPRPTSHARPPSTSLIYDSDLA
		ļ			I	VSYTDLDNLFNSDEDELTPGSKRSANGSDDK
						ASCKESKTGNLDPLSCISTADLHKMYPTPPSL
						EQHIMGFSPMNMNNKEYGSMDTTPGGTVLE
				-	ł	GNSSSIGAQFKIEVDEGFCSPKPSEIKDFSYVY
			i .			KPENCQILVGCSMFAPLKTLPSQYLPLIKLPEE
						CIYRQSWTVGKLELLSSGPSMPFIKEGDGSNM
						DQEYGTAYTPQTHTSCGMPPSSAPPSNSGAGI
}						LPSPSTPRFPTPRTPRTPRTPRGAGGPASAQGS
						VKYENSDLYSPASTPSTCRPLNSVEPATVPSIP
			[			EAHSLYVNLILSESVMNLFKDCNSDSCCICVC
1	,		]			NMNIKGADVGVYIPDPTQEAQYRCTCGFSAV
			ļ l			MNRKFGNNSGLFFEDELDIIGRNTDCGKEAE
			1			KRFEALRATSAEHVNGGLKESEKLSDDLILLL QDQCTNLFSPFGAADQDPFPKSGVISNWVRV
]				•		EERDCCNDCYLALEHGRQFMDNMSGGKVDE
			}			ALVKSSCLHPWSKRNDVSMQCSQDILRMLLS
}						LQPVLQDAIQKKRTVRPWGVQGPLTWQQFH
						KMAGRGSYGTDESPEPLPIPTFLLGYDYDYLV
						LSPFALPYWERLMLEPYGSQRDIAYVVLCPE
				-		NEALLNGAKSFFRDLTAIYESCRLGQHRPVSR
			[ ]			LLTDGIMRVGSTASKKLSEKLVAEWFSQAAD
			}			GNNEAFSKLKLYAQVCRYDLGPYLASLPLDS
)	1			0.0		SLLSQPNLVAPTSQSLITPPQMTNTGNANTPS
						ATLASAASSTMTVTSGVAISTSVATANSTLTT
						ASTSSSSSNLNSGVSSNKLPSFPFFGSMNSNA
i					ĺ	AGSMSTQANTVQSGQLGGQQTSALQTAGISG
						ESSSLPTQPHPDVSESTMDRDKVGIPTDGDSH AVTYPPAIVVYIIDPFTYENTDESTNSSSVWTL
						GLLRCFLEMVQTLPPHIKSTVSVQIIPCQYLLQ
						PVKHEDREIYPQHLKSLAFSAFTQCRRPLPTS
	ļ				l	TNVKTLTGFGPGLAMETALRSPDRPECIRLYA
						PPFILAPVKDKQTELGETFGEAGQKYNVLFV
						GYCLSHDQRWILASCTDLYGELLETCIINIDVP
	. 1			l		NRARRKKSSARKFGLQKLWEWCLGLVQMSS
	ĺ			l		LPWRVVIGRLGRIGHGELKDWSCLLSRRNLQ
	ļ	٠		1		SLSKRLKDMCRMCGISAADSPSILSACLVAM
						EPQGSFVIMPDSVSTGSVFGRSTTLNMQTSQL
		j			J	NTPQDTSCTHILVFPTSASVQVASATYTTENL
.		ì			1	DLAFNPNNDGADGMGIFDLLDTGDDLDPDII
	ŀ				I	NILPASPTGSPVHSPGSHYPHGGDAGKGQSTD
						RLLSTEPHEEVPNILQQPLALGYFVSTAKAGP
	· · · · · · · · · · · · · · · · · · ·					LPDWFWSACPQAQYQCPLFLKASLHLHVPSV
	· 1				1	QSDELLHSKHSHPLDSNQTSDVLRFVLEQYN ALSWLTCDPATQDRRSCLPIHFVVLNQLYNFI
1	1	- 1	ł	1	į.	MNML
789	2139	A	6359	1	2002	TGTLTEDGLDVMGVVPLKGQAFLPLVPEPRR
				-		LPVGPLLRALATCHALSRLQDTPVGDPMDLK
	j			ļ	l	MVESTGWVLEEPAADSAFGTQVLAVMRPP
		į		I	_ , I	LWEPQLQAMEEPPVPVSVLHRFPFSSALQRM
- 1		ı		Ì	·	SVVVAWPGATQPEAYVKGSPELVAGLCNPET
	j			1	i	VPTDFAQMLQSYTAAGYRVVALASKPLPSVP
		ļ	1	ļ	i	SLEAAQQLTRDTVEGDLSLLGLLVMRNLLKP
[			1	1	l	QTTPVIQALRRTRIRAVMVTGDNLQTAVTVA
		ļ	1	l	i	RGCGMVAPQEHLIIVHATHPERGOPASLEFLP
	į	[	1	l	ŀ	MESPTAVNGVKDPDQAASYTVEPDPRSRHLA
		l	i		ŀ	LSGPTFGIIVKHFPKLLPKVLVQGTVFARMAP
.			- 1	-		EQKTELVCELQKLQYCVGMCGDGANDCGAL

SEQ ID	SEQ ID	Met	CPA	Predicted	Dandista d and	Amino cald names (A. Alasta C.C.)
			SEQ	i .	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	Į .	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
dello			714	amino acid		Q-Giutannie, K-Arginine, S-Serine,
1					of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	į	/=possible nucleotide deletion, \=possible
	1		1	sequence	1	nucleotide insertion
	<del>                                     </del>		<del>                                     </del>	<u> </u>	<del> </del>	VIREGRCSLDTSFSVFKYMALYSLTQFISVLIL
ł	l	l	l		ì	AIVEOVCORDI 2L2ALK I MAT I 2TI OLIZATIT
İ						YTINTNLGDLQFLAIDLVITTTVAVLMSRTGP
1			l .	1		ALVLGRVRPPGALLSVPVLSSLLLQMVLVTG
1						VQLGGYFLTLAQPWFVPLNRTVAAPDNLPNY
1 .	1			1	f	ENTVVFSLSSFQYLILAAAVSKGAPFR\RPLTN
	1	1			!	NVPFLLASAL*SSVLVVLVLSPGLLHGPLALR
i	İ			ì	l	
1	1			ł	ļ	NTTDTGFKLLLVGLVTLNFVGGLHAGERARP
1	1		i	ł	ļ	VPPRLPAPPPAQAG\SKKRFKQLERELAEQPW
					i	PPLPAGPLR
790	2140	Α	6380	76	1059	SSAGSARKLQVMALAARLWRLLPFRRGAAP
i		l	I	l -		GSRLPAGTSGSRGHCGPCRFRGFEVMGNPGT
1		l	l	1		
	1	l		1		FKRGLLLSALSYLGFETYQVISQAAVVHATA
1.		1				KVEEILEQADYLYESGETEKLYQLLTQYKESE
I			ľ			DAELLWRLARASRDVAQLSRTSEEEKKLLVY
1	ì	l	1	l		EALEYAKRA/L/EKNESSFASHKWYAICLSDV
1		1		l		GDYEGIKAKIANAYIIKEHFEKAIELNPKDATS
1	1	1				IHLMGIWCYTFAEMPWYORRIA*NACLOLPP
1		1				
		1				*FPPYEKALG\YFHRAEQVDPNFYSKNLLLLG
1	l	l				KTYLKLHNKKLAAFWLMKAKDYPAHTEED
		1				KQIQTEAAQLLTSFSEKN
791	2141	Α	6434	3	1460	IALLIVDGLAWDDQGGLALLHISPSKLIL+QDS
				_		SGMS/YVMVRCTITRAFFKSLLCHICOYSIGPO
	]	ĺ				
Į.						*VT\CPGQDACKE*KSTAN*GG*RE**PQVLFF
1	1					AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ
1						RLQEQRQQQSGEAEALARVYSSSISNGLSNLN
						NETSGTYANGSVIDLPKSEGYYNVVSGQPSP
ł						DQSGLDMT\GIKQIKQEPIYDLTSVPNLFTY\SS
<b>1</b> .						FNN/GQLAPGIT\MTEIDRIAQNIIKSHLETCQY
	1					TAGET TO AMOTITE TERM MORE THE 1CO
						TMEELHQLAWQTHTYEEIKAYQSKSREALW
						QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ
l'						ILLLKSGCLEVVLVRMCRAFNPLNNTVLFEG
	<b>[·</b>					KYGGMQMFKALGSDDLVNEAFDFAKNLCSL
1						QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ
						EKIYFALOHVIQKNHLDDETLAKLIAKIPTITA
1 .					1	VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF
			Ll			NPDCATACK
792	2142	Α	6440	92	781	SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV
				İ		TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL
]	]					MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT
[	į					LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA
1	i		! !			LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH
			{		-	ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH
				·	l	GAVVNESHHDALVEDIFDKEDEDKDGFISAR
				ļ		EFTYKHDEL
793	2143	Α	6446	3201	152	PRLKRLVVTEEDGGARPEALGKIAPRTPAELG
ا "' ا	2173	Λ.	J-7-70	J201	132	
				}		ARADQELVTALMCDLRRPAAGGMMDLAYV
l l				ł		CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM
Į į			1			DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH
1 1			. 1		ļ	EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS
1 1	į į			ľ		MALCHANDOUTPOSTAGE INCOME.
					1	MGVSTLAINS WE\SSVGSLIVEGGPHLWALSI
			{			WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS
j			1		1	P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLO\P
1						SGQVL\TST\ESLCRLRARVALADIAFTGGGNI
					ł	7
<b>!</b>			]		1	VVATADGSSA\SPVQFYKVCV\$VVSEKCRIDT
			<u> </u>	ı		DILPSLFMRCTTDLNRKDKFPAITHLKFLARD
[			'	1	.4	MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI
			ŀ			FQQISPVVGDKQPTILKWRILSATNDLDRVSA
}		ļ	l	.	į	VALPKLPISLTNIDLKVASDTQFYPGLGLAL
j (		- 1	, , <b>, l</b>	ł		AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD
			<b>' I</b>			
	L					EPAMKRPRTAGPAVHLKAMQLSWTSLALVG

WO 01/57188

PCT/US01/03800

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	seq-		USSN 09/496	location	to last amino	I=Isoleucine, K=Lysine, L=Leucine,
uence	ualac		914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
		1	7.4	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
-				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1			peptide		/=possible nucleotide deletion, \=possible
	<u> </u>		1	sequence		nucleotide insertion
						IDSHGKLSV\LRLSPSMGHPLEVGLALRHLLFL
ł	1	1		1	1	LEYCMVTGYDWWDILLHVQPSMVQSLVEKL
						HEEYTRQTAALQQVLSTRILAMKASLCKLSP
						CTVTRVCDYHTKLFLIAISSTLKSLLRPHFLNT
						PDKSPGDRLTEICTKITDVDIDKVMINLKTEEF VLDMNTLQALQQLLQWVGDFVLYLLASLPN
1	1	1	[	[	ł	QPCPTSEPCPTSEPSPTSEPSPTSEPSP*SLC\G
						SLLRPGHSFLRDGTSLGMLRELMVVIRIWGLL
ł	-	1	1			KPSCLPVYTATSDTQDSMSLLFRLLTKLWICC
		*		<b>!</b>		RDEGPASEPDEALVDECCLLPSQLLIPSLDWL
1		l				PASDGLVSRLQPKQPLRLQFGRAPTLPGSAAT
	1	l				LQLDGLARAPGQPKIDHLRRLHLGACPTEEC
				[	[	KACTRCGCVTMLKSPNRTTAVKQWEQRWIK
1				İ		NC/LVRWALVAGAPQLPLSPAAPQLLLSYPSA APEPGCCKSHRSPWTLLGAVNLSPPCRAVEG
1		l	}			RGPDACVTSRASEEAPAFVQLGPQSTHHSPRT
				,		PRSLDHLHPEDRP
794	2144	Α	6490	418	585	NGDKADLENESCRAQVLMPVVPALWEAEGG
						GSIEPRDLRLQ*AVITPL\TPAWVTO
795	2145	A	6499	395	1027	KLLWLPPHSEQKRSPLYHPQGPSGTTPSAP\FS
						SHSPPPSLLQAVPSIAAFLRTHGHISASGPLRMP
						FPH/H*NAFLLVFPGQRSQLTS/PSHYLCREVFP
		Ĭ				DHHHHLCRLSLESSPLFHHRVLFCVPKQNVN STRAQIFCLFVHIVGCRCINTFPLHLFRLHLWL
				`		HFLQIPLCKKNKSVKLGKTVVGRGCQSAAGS
		[ [				DTRVRAAVGAPGLPVEPLV
796	2146	A	6503	68	936	HSALLTHSSFCVFTLCQDFFTYSSMSEEVTYA
*					,	DLQFQNSSEMEKIPEIGKFGEKAPPAPSHVWR
						PAALFLTLLCLLLLIGLGVLASMFHVTLKIEM
						KKMNKLQNISEELQRNISLQLMSNMNISNKIR
						NLSTTLQTIATKLCRELYSKEQEHKCKPCPRR WIWHKDSCYFLSDDVQTWQESKMACAAQN
						ASLLKINNKNALEFIKSQSRSYDYWLGLSPEE
						DS/YSWYESG*YNQ\PSAWVIRNAPDL\N\MY
						CGYINRLYVQYYHCTYKQRMICEKMANPVQ
						LGSTYFREA
797	2147	A	6507	1	881	PGSTHASARSQVPRSAGEAAPHSRRPPGLLPH
					·	APRAASAQLEERMRDPHPGMTLQEGDCRGS
						QTVSLTMGTADSDEMAPEAPQHTHIDVHIHQ
						ESALAKLLLTCCSALRPRATQARGSSRLLVAS WVMQIVLGILSAVLGGFFYIRDYTLLVTSGA
	}		+	· 1		AIWTGAVAVLAGAAAFIYEKRGGTYWALLR
]		']				TLLALAAFSTAIAALKLWNEDFRYGYSYYNS
· ]			J	l		ACRISSSSDWNTPAPTQSPEEVRRLHLCTSFM
			[	1	Í	DMLKALFRTLQAMLLGVWILLLASLTPLWL
		ì	- 1		1	/SL/RGECSQPKG*VPKKRDQKEMLEVSGI*PG
ŀ	.	ł	ł		ļ	STHASARSQVPRSAGEAAPHSRRPPGLLPHAP
			1	1	ļ	RAASAQLEERMRDPHPGMTLQEGDCRGSQT
			1			VSLTMGTADSDEMAPEAPQHTHIDVHIHQÈS
		ļ		]		ALAKLLITCCSALRPRATQARGSSRLLVASW VMQIVLGILSAVLGGFFYIRDYTLLVTSGAAI
	ĺ	- 1	ſ	ĺ	· 1	WTGAVAVLAGAAAFIYEKRGGTYWALLRTL
l	1	ı		ļ		LALAAFSTAIAALKLWNEDFRYGYSYYNSAC
ł	l	ł	1	ļ	ì	RISSSDWNTPAPTQSPEEVRRLHLCTSFMDM
•			Ī	i		LKALFRTLQAMLLGVWILLLLASLTPLWLYC
700			(865			WRMFPTKGVSP
798	2148	A	6528	912	2287	VPNYLPSVSSAIGGEVPQRYVWRFCIGLHSAP
	j		·	1		RFLVAFAYWNHYLSCTSPCSCYRPLCRLNFG
					<u></u>	LNVVENLALLVLTYVSSSEDF/TWVPG*GRSG

WO 01/57188

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted and	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	соптевропой	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Londo	1		1 224	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ł	1	i	ı	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			l		sequence	
			1	peptide		/=possible nucleotide deletion, \=possible
			<b>-</b> .	sequence		nucleotide insertion
Į.						EVFPEGTGLPLPHSDLPTSWCGHSLQCGSQSS
İ					1	FPPAIHENAFIVFIASSLGHMLLTCILWRLTKK
	ł					HTVSQE\DGLSLAGAPRQPRRKSRTSVLRIRV
		İ	i			MVRWELSSNGNPGRGVLGLGLGLGNKLRVV
i		İ			1	GQNLGL*HCVWVVWETGE*KRWRLQMGIE*
i l	j		i		i	GVASRRQ*VRNSVRGLVCHNSSAPPMYMGFF
[	ľ		<b>!</b>			SPTVFGGGVGG*LHVTFILHPPEVEAAGIPLLL
	ļ		<b>i</b> 1			GPSLPQRQGREHIVVILAAPACAPFHDR*WEP
			[			REIRPSP*ELGLRGEPTLSYPASCRVIROPIP*D
	i	Į .	1		ļ	l <b>7</b>
		1	!		İ	RKSYSWKQRLFIINFISFFSALAVYFRHNMYC
	ĺ	ĺ				EAGVYTIFAILEYTVVLTNMAFHMTAWWDF
700	0140	-	(600		054	GNKELLITSQPEEKRF
799	2149	A	6529	1 .	874	FFFFQRINFIEHSGSVSLLALACDLGWCEDWS
	ŀ					CCLVQGGGDLVDVVQTNHGEDEAGGDTDSV
	Ì			ĺ		DEARCKESQQEAQENLREDLCLESFAKDKIL
						QUEGSEREHEETRTKQAALDGEPLGGGQLTA
			1			VHLHPSKEQQGQEGGERQRGARTHHWRGW
	ļ		l i			EKGRRVRLRPPSGKLRADQPVRKLGGPTPS/T
	ł					ELPGLQPHAPTPHTA/PATPTYSPAPDTPNPPV
	1					RWKCPLPVEPRTRQLCRERTRKACPPKPRPPL
			i i			GLPGDPTGPVTHHAPPVSPTGASGOERRAEP
				İ		GAVSYAHASATK
800	2150	Α	6544	2	662	SAQRWAAVAGRWGCRLLALLLLVPGPGGAS
	2.50	**	0544	_	002	EITFELPDNAKQCFYEDIAQGTKCTLEFQVITG
				·		GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF
	i					TASKNGTYKFCFSNE\FSTFTHKTVYFDFQVG
					_ :	EXTHLCFLVR/DRVSALTQMESACVSIHEALKS
						VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV
003	0151					GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS
801	2151	A	6556	1	1319	TPCMECIKGEGLREPQNLSGSQREPQTEGSM
	· ·	i				DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF
						KRIFLKRMPSIRESLKERGVDMARLGPEWSQP
				1		MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP
			1			PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY
						HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL
				i		SQDITVGGITVTQMFGEVTEMPALPFMLAEF
				ŀ		DGVVGMGFIEQAIGRVTPIFDNIISOGVLKED
				,	-	VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE
		I				GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE
		I	1			DOOR AT TIMES A STREET CONTRACTOR A TO A TO A STREET
		ł			·	KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT
		1				SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL
		1		00		· · · ·
802	2152		6867	12	6147	VALGATF\IRKFYTEFDRGNNPHGFALAR
0U2	2152	A	6567	13	6147	MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL
'		. [	1	l		LAVVVLLALPVAWGQCNAPEW\LPFARPTNL
					Ì	TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS
		l		ļ		VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG
		ļ	l			IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW
		ļ		1		DNETPICDRIPCGLPPTITNGDFISTNRENFHY
		1				GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND
				Ī		DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
. [		į	1	1		NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA
		1	I	ľ		LNK WEPELPSCSR V COPPPDVLHAERT ORDK
ł		l	ł	ł	ł	•
	ĺ		ļ	l	ļ	DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG
1	ľ		ļ	l	į	DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV
						NLQLGAKVDFVCDEGFQLKGSSASYCVLAG
1		I	i	l		-
[			j			MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP
						-
						MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP
						MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR

SEQ ID NO: of nucl- cotide seq-	SEQ ID NO: of peptide	Met hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
nucl- cotide		hod	I ID NO.			
cotide	l mantida		I III NO:	beginning	nuclcotide	D=Aspartic Acid, E=Glutamic Acid,
	pepude		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cen.	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
344	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		l	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
}		ļ		peptide	Soquetto	/=possible nucleotide deletion, \=possible
1	i			sequence		
	<del></del>	<del></del>	<del></del>	sequence		nucleotide insertion
1	l		1			SITCLDNLVWSSPKDVCKRKSCKTPPDPVNG
			<u>.</u>			MVHVITDIQVGSRINYSCTTGHRLIGHSSAECI
1			į į			LSGNAAHWSTKPPICQRIPCGLPPTIANGDFIS
		İ				TNRENFHYGSVVTYRCNPGSGGRKVFELVGE
		İ				PSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
	ŀ	İ				ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGP
, ,		J	j j			RRVKCQALNKWEPELPSCSRVCQPPPDVLHA
					•	ERTORDKDNFSPGQEVFYSCEPGYDLRGAAS
	i		i I			MRCTPQGDWSPAAPTCEVKSCDDFMGQLLN
i			ł l			
]						GRVLFPVNLQLGAKVDFVCDEGFQLKGSSAS
		1				YCVLAGMESLWNSSVPVCEQIFCPSPPVIPNG
, 1			)			RHTGKPLEVFPFGKAVNYTCDPHPDRGTSFD
						LIGESTIRCTSDPQGNGVWSSPAPRCGILGHC
		}				QAPDHFLFAKLKTQTNASDFPIGTSLKYECRP
1 1						<b>EYYGRPFSITCLDNLVWSSPKDVCKRKSCKTP</b>
						PDPVNGMVHVITDIQVGSRINYSCTTGHRLIG
		1				HSSAECILSGNTAHWSTKPPICQRIPCGLPPTI
1 1			1			ANGDFISTNRENFHYGSVVTYRCNLGSRGRK
1 1				1		VFELVGEPSIYCTSNDDQVGIWSGPAPOCIIPN
<b>i</b> I						KCTPPNVENGILVSDNRSLFSLNEVVEFRCOP
1 1				ļ		GFVMKGPRRVKCQALNKWEPELPSCSRVCQ
1 !				ł		PPPEILHGEHTPSHQDNFSPGQEVFYSCEPGY
1 1						DI DOA Y SI ROLLO COMODO VIDO VIDO CONDE
i 1			ĺ			DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF
i l						LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL
1 1				,		KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPN
1 1						PPAILNGRHTGTPSGDIPYGKEISYTCDPHPDR
i l	i					GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE
i l			f		.	LSVRAGHCKTPEQFPFASPTIPINDFEFPVGTS
1			ı		İ	LNYECRPGYFGKMFSISCLENLVWSSVEDNC
i '						RRKSCGPPPEPFNGMVHINTDTQFGSTVNYSC
1 1			l	j		NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII
i [			}			SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH
i l	ŀ		1			TGPDGEQLFELVGERSIYCTSKDDQVGVWSS
i l	1		١ .	ı		PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI
i l	1			ŀ		IRFRCQPGFVMVGSHTVOCOTNGRWGPKLPH
1	İ		- 1	į		CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY
. 1	j			ľ		CCEDENDI DO Y TO TICEDO COMO DE LA COLLA C
i I						SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV
; I	ł	l	1	ł	}	KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC
, 1		ı	- 1		1	DEGFRLKGRSASHCVLAGMKALWNSSVPVC
		- 1	- 1			EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT
, I			ł			CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS
, I		.	- 1	ŀ	l	SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL
, I			- 1	ł	1	YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS
, J	}	Į	1	J	j	OLDHYCKEVNCSFPLFMNGISKELEMKKVYH
i	}			ļ	i	YGDYVTLKCEDGYTLEGSPWSQCQADDRWD
	ļ	ļ	1	i		PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI
	1	l		l	l	ILKHRKGNNAHENPKEVAIHLHSOGGSSVHP
,	1			l		
803	2153	<del></del>	6574	<del>,</del>	2022	RTLQTNEENSRVLP
303	4133	A	03/4	2	3233	HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL
	i	ĺ	1	ſ		LLLPPLLLLRG\SHAGNLTVAVVLPLANTSY
	ļ	I	ļ			PWSWA\RVGPAVELALAQVKARPDLLPGWT
	I		1			VRTVLGSSENALGVCSDTAAPLAAVDLKWE
I	j	ļ	1		l	HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL
- 1		l	1		[	TAGAPALGFGVKDEYALTTRAGPSYAKLGDF
- 1		j	ļ		*	VAALHRRLGWERQALMLYAYRPGDEEHCFF
i		Ì			`	
j		ľ	J		. !	LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT
ŀ		I				RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA
l		J		-		GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW
l		- 1		1		ERGDGQDVSARQAFQAAKIITYKDPDNPEYL
	l l	- 1				EFLKQLKHLAYEQFNFTMEDGLVNTIPASFH

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	bod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Ghttamic Acid,
nucl-	peptide		in	nucleotide	location	
cotide	seq-	1	USSN	location		F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence	i	4		corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	uence	ľ	09/496	correspondi	to last amino	M=Methlonine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	l	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	i	i	ł	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		Ī	1	peptide		/=possible nucleotide deletion, \=possible
				sequence	L	nucleotide insertion
1	ł	1	ł			DGLLLYIQAVTETLAHGGTVTDGENITQRMW
		i		ŀ		NRSFQGVTGYLKIDSSGDRETDFSLWDMDPE
	ĺ	1			ì	NGAFRVVLNYNGTSQELVAVSGRKLNWPLG
Į.		t		J	j	YPPPDIPKCGFDNEDPACNQDHLSTLEVLALV
1	İ	l				GSLSLLGILIVSFFIYRKMOLEKELASELWRVR
1		ŀ				WEDVEPSSLERHLRSAGSRLTLSGRGSNYGSL
J i	1		]		i	LTTEGQFQVFAKTAYYKGNLVAVKRVNRKR
			1			IELTRKVLFELKHMRDVQNEHLTRFVGACTD
1			1			PPNICILTEYCPRGSLQDILENESITLDWMFRY
			i			SLTNDIVKGMLFLHNGAICSHGNLKSSNCVV
			1 1			
1			1 1			DGRFVLKITDYGLESFRDLDPEQGHTVYAKK
						LWTAPELLRMASPPVRGSQAGDVYSFGIILQE
1				i		IALRSGVFHVEGLDLSPKEIIERVTRGEQPPFR
						PSLALQSHLEELGLLMQRCWAEDPQERPPFQ
				·		QIRLTLRKFNRENSSNILDNLLSRMEQYANNL
j i					·	EELVEERTQAYLEEKRKAEALLYQILPHSVAE
1 1			i I			QLKRGETVQAEAFDSVTIYFSDIVGFTALSAE
			[ [			STPMQVVILLNDLYTCFDAVIDNFDVYKVET
						IGDAYMVVSGLPVRNGRLHACEVARMALAL
1 1			1 1	'		LDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGV
1 1						VGLKMPRYCLFGDTVNTASRMESNGEAL\KI
				·		HLSS/BTKAVL/EEFGGFELELRGDVEMKGKG
804	2154	A	6585		3837	KYRTYWLLGERGSSTRG
1007	2.57	^	[ 6365 ]	<i>-</i> 1	3037	DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV
1 1						
	1					MSERVSGLAGSIYREFERLIVRYDEEVVKELIP
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE
				·		LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT
	:					LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP
·						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLTTQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLBEKNRELBEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLTTQYEREKALRKHAEEKFIEPEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLBEKNRELBEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWOFFSRLFSSSNITK
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGLEAVKQAKLKLBEKNRELBEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWOFFSRLFSSSSNITIK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLBEKNRELBEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNITIK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPOKYKOVTNG
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLBEKNRELBEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFPSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEPEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLTTQYEREKALRKHAEEKFIEPEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLBEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLTTQYEREKALRKHAEEKFIEPEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLBEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKWVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLBEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFPSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ TEGSKQRSASQSSLDKLDQELKEQQKELKNQ TEGSKQRSASQSSLDKLDQELKEQQKELKNQ TEGSKQRSASQSSLDKLDGELKEQGKELKNQ TEGSKQRSASQSSLDKLDGELKEQGKELKNQ TEGSKQRSASQSSLDKLDGELKEQGKELKNQ TEGSKQRSASQSSLDKLDGELKEQGKELKNQ TEGSKQRSASQSSLDKLDGELKEQGKELKNQ TEGSKQRSASQSSLDKLDGELKEQGKELKNQ TEGSKQRSASQSSLDKLDGELKEQGKELKNQ TEGSKQRSASQSSLDKLDGELKEQGKELKNQ TEGSKQRSASGSSLDKLDGELKEQGKELKNQ TEGSKQRSASQSSLDKLDGELKEQGKELKNQ TEGSKQRSASQSSLDKLDGELKEQGKELKNQ TEGSKQRSASGSSLDKLDGELKEQGKELKNQ TEGSKQRSASGSSLDKLDGELKEQGKELKNQ TEGSKQRSASGSSLDKLDGELKEQGKELKNQ TEGSKQRSASGSSLDKLDGELKEQGKELKNQ TEGSKQRSASGSSLDKLDGELKEQGKELKNQ TEGSKQRSASGSSLDKLDGELKEQGKELKNQ TEGSKQRSASGSSLDKLDGELKEQGKELKNQ TEGSKQRSASGSSLDKLDGELKEQGKELKNQ TEGSKQRSASGSSLDKLDGELKEQGKELKNQ TEGSKQRSASGSSLDKLDGELKEQGREDLSESG TVCNSHVLCIASVPGARETDYPAGEDLSESG TVCNSHVLCIASVPGARETDYPAGEDLSESG TVCNSHVLCIASVPGARETDYPAGEDLSESG TVCNSHVLCIASVPGARETDYPAGEDLSESG TVCNSHVLCIASVPGARETDYPAGEDLSESG TVCNSHVLCIASVPGARETDYPAGEDLSESG
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEELIRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWOFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENYPTABEVATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGVQIPEDLSPVYQSSND
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEPEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTABEVATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGVQPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLTTQYEREKALRKHAEEKFIEPEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLBEKNRELBEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNILKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLTTQYEREKALRKHAEEKFIEPEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLBEKNRELBEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTABEVATEATEGNAGSAEDTVDIS QTGYYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGGW
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLBEKNRELBEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNITK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSBETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTABEVATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLBEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFPSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEVATEATEGNAGSAEDTVDIS QTGYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHILLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESOVRO
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ CELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEVATEATEGNAGSAEDTVDIS QTGYYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYOHLO
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLBEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFPSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEVATEATEGNAGSAEDTVDIS QTGYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHILLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESOVRO

SEQ ID NO: of nucl- eotide seq-	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496	Predicted beginning nucleotide location correspondi	Predicted end nucleotide location corresponding to last amino	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid residue of peptide sequence	acid residue of peptide sequence	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *-Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						NRLWVGTGNGVIISIPLTETVILHQGRLLGLR ANKTSGVPGNRPGSVIRVYGDENSDKVTPGT FIPYCSMAHAQLCFHGHRDAVKFFVAVPGQV ISPQSSSSGTDLTGDKGRGHLHRSLVVRRP
805	2155	A	6605	469	2602	FGRILWGTAFKSWKMKAPIPHLII.LYATFTQ SLKVVTKRGSADGCTDWSIDIKKYQVLVGEP VRIKCALFYGYIRTNYSLAQSAGLSLMWYKS SGPGDFEEPIAFDGSRMSKEEDSIWFRPTLLQ DSGLYACVIRNSTYCMKVSISLTVGENDTGL CYNSKMKYFEKAELSKSKEISCRDIEDFLLPT
			·			REPEILWYKECRTKTWRPSIVFKRDTILLIREV REDDIGNYTCELKYGGFVVRRTTELTVTAPL TDKPPKLLYPMESKLTIQETQLGDSANLTCRA FFGYSGDVSPLIYWMKGEKFIEDLDENRVWE SDI\KILKEHLGEQEVSISLIVDSVEEGDLGNYS
						CYVENGNGRRHASVLLHKRELMYTVELAGG LGAILLLLVCLVTIYKCYKIEIMLFYRNHFGA EELDGDNKDYDAYLSYTKVDPDQWNQETGE EERFALEILPDMLEKHYGYKLFIPDRDLIPTGT YIEDVARCVDQSKRLIIVMTPNYVVRGWSIF ELETRLRNMLVTGEIKVILIECSELRGIMNYQE VEALKHTIKLLTVIKWHGPKCNKLNSKFWKR LQYEMPFKRIEPITHEQALDVSEQGPFGELQT
						VSAISMAAATSTALATAHPDLRSTFHNTYHS QMRQKHYYRSYEYDVPPTGTLPLTSIGNQHT YCNIPMTLINGQRPQTKSSREQNPDEAHTNSA ILPLLPRETSISSVIW
806	2156	A	6614	3	1584	NSARGGVGVRGARAMATVQEKAAALNLSAL HSPAHRPPGFSVAQKPFGATYVWSSIINTLQT QVEVKKRRHRLKRHNDCFVGSEAVDVIFSHIL IQNKYFGDVDIPRAKVVRVCQALMDYKVFE AVPTKVFGKDKKPTFEDSSCSLYRFTTIPNQD SQLGKENKLYSPARYADALFKSSDIRSASLED
					·	LWENLSLKPANSPHVNISTTLSPQVINEVWQE ETIGRLLQLVDLPLLDSLLKQQEAVPKIPQPK RQSTMVNSSNYLDRGILKAYSDSQEDEWLSA AIDCLEYLPDQMVVEISRSFPEQPDRTDLVKE LLFDAIGRYYSSREPLLNHLSDVHNGIAELLV
				·		NGKTEIALEATQLLLKILLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYQQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGOFYKCHPDIFIEHFGD
807	2157	A	6615	4198	2094	FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG
				į		GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQLAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR BTGMPDDSSCGMARSCOVEALBY
						LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL NLSRNCLRTFEARRLGSLPCLMILDLSHNALE TLELGARALG\SLRTLLLQGNALRDLPPYTFA NLASLQRLNLQGNRVSPCGGPDEPGP\SGCV\ AFSGITSLRSLSLVDNEIELLRAGAFLHTPLTE

052		157	1 050	18	Tn 4	
SEQ ID		Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	į.	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ì	1	1	ı	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1		J	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i	i		1	peptide		/=possible nucleotide deletion, \-possible
L		L.	1	sequence		nucleotide insertion
	1	1				LDLSSNPGLEVATGALGGLEASLEVLALQGN
l	i	Ī				GLMVLQVDLPCFICLKRLNLAENRLSHLPAW
ı			1	}		TQAVSLEVLDLRNNSFSLLPGSAMGGLETSLR
1		1	.		i	RLYLQGNPLSCCGNGWLAAQLHQGRVDVDA
I	1	1	1	I	ļ	TQDLICRFSSQEEVSLSHVRPEDCEKGGLKNI
I		1	1			NLIILTFILVSAILLTTLAACCCVRRQKFNQQ
	1	J				YKA
808	2158	A	6619	153	1852	FKALSQYIYTNTHLEREAAFEVAILLRRMEEG
		1 .	1		1002	ARHRNNTEKKHPGGGESDASPEAGSGGGGV
		1	ĺ	F .	1	ALKKEIGLVSACGIIVGNIIGSGIFVSPKGVLEN
	I	i				AGSVGLALIVWIVTGFITVVGALCYAELGVNI
•	1	ı	1		i	PKSGGDYFYVKDIFGGLAGFLRLWIAVLVTYP
1		İ			χ.	TNQAVIALTESNYVLQPLEPTCEPPESGLELLA
1	1	1	1			AICLLLLTWVNCSSVRWATRVQDIFTAGKLL
1	1	1	1		l '	ALAL HIM COUNTY OF THE TOTAL T
1	[	1				ALALIIMGIVQICKGEYFWLEPKNAFENFQEP
	1	l				DIGLVALAFLQGSFAYGGWNFLNY\VTEELV
1	1	l			1	DP\YKNL\PRAIFISIP\LVTFVYVFANV/ALYVT
1			j l			AMSPQELLASINAVAVTFGEKLLGVMAWIM
						PISVALSTFGGVNGSLFTSSRLFFAGAREGHLP
	[		}			SVLAMIHVKRCTPIPALLFTCISTLLMLVTSD
	į	i			'	MYTLINYVGFINYLFYGVTVAGQIVLRWKKP
	1					DIPRPIKINLLFPITYLLFWAFLLVFSLWSEPVV
1	1	1	ŀ			CGIGLAIMLTGVPVYFLGVYWQHKPKCFSDFI
	İ	]				ELLTLVSQKMCVVVYPEVERGSGTEEANED
						MEEQQQPMYQPTPTKDKDVAGQPQP
809	2159	A	6621	1041	223	QDSRKMLPSTSVNSLVQGNGVLNSRDAARH
]						TAGAKRYKYLRRLFRFRQMDFEFAAWQMLY
1			! [			LFTSPQRVYRNFHYRKQTKDQWARDDPAFL
I			1 1			VLLSIWLCVSTIGFGFVLDMGFFETIKLLLWV
[			]			VLIDCVGVGLLIATLMWFISNKYLVKRQSRD
ĺ			f 1			YDVEWGYAFDVHLNAFYPLLVILHFIQLFFIN
<b>[</b>	1 1		1 1	ĺ		HVILTDTFIGYLVGNTLWLVAVGYYIYVTFL
	1 1		<b>!</b> !			GYSVGLLFFS\ALPFLKNTVILLYPFAPLILLYG
						LSLALGWNFTHTLCSFYKYRVK
810	2160	A	6623	160	822	SPASGHCRLNGAAVAMFGCLVAGRLVQTAA
			1 1	1		QQVAEDKFVFDLPDYESINHVVVFMLGTIPFP
				j		EGMGGSVYFSYPDSNGMPVWQLLGFVTNGK
				1		PSAIFKISGLKSGEGSQHPFGAMNIVRTPSVAQ
					İ	IGISVELLDSMAQQTPVGNAAVSSVDSFTOFT
				ĺ		QKMLDNFYNFASSFAVSQ/VPDDTQ/RPSEMF
				l		IPANVVLKWYENFORRTSTEPSLLENIIWIKIN
				1		F
811	2161	Α	6627	18	3367	LEGSLNTERAKYYLTITMPHFTVTKVEDPEEG
		l		1		AAASISQEPSLADIKARIQDSDEPDLSQNSITG
				l		EHSQLLDDGHKKARNAYLNNSNYEEGDEYF
	ļ. <b>I</b>	- 1		}	l l	DKNLALFEEEMDTRPKVSSLLNRMANYTNLT
		1	J	l		QGAKEHEEAENITEGKKKPTKTPOMGTFMG
				1		VYLPCLQNIFGVILFLRLTWVVGTAGVLQAF
	ŀ	ľ	·		i	AIVLICCCCTMLTAISMSAIATNGVVPAGGSY
	l	- 1		1	Į.	
		í		ŀ		FMISRALGPEFGGAVGLCFYLGTTFAAAMYIL
0.0	ſ	I	ŀ		1	GAIEIFLVYIVPRAAIFHSDDALKESAAMLNN
			İ		Ì	MRVYGTAFLVLMVLVVFIGVRYVNKFASLFL
	l l	1	l	İ	j	ACVIVSILAIYAGAIKSSFAPPHFPVCMLGNRT
		1	ļ		. I	LSSRHIDVCSKTKEINNMTVPSKLWGFFCNSS
ļ	ŀ	ľ	1		ļ	QFFNATCDEYFVHNNVTSIQGIPGLASGITEN
ł	ł	- 1	ł		. 1	LWSNYLPKGEIIEKPSAKSSDVLGSLNHEYVL
		- 1				VDITTSFTLLVGIFFPSVTGIMAGSNRSGDLKD
			1	I		AQKSIPIGTILAILTTSFVYLSNVVLFGACIEGV
1		- 1		1		AN DESCRIPTION AND AND AND AND AND AND AND AND AND AN
			1	- 1		VLKUKFGDA VKGNL V VGILSWPSPW VIVIGS I
	1	ļ	}		;	VLRDKFGDAVKGNLVVGTLSWPSPWVIVIGS FFSTCGAGLQSLTGAPRLLQAIAKDNIIPFLRV

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	i			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \-possible
<b> </b>	ļ			sequence		nucleotide insertion
	1					FGHSKANGEPTWALLLTAAIAELGILIASLDL
J	]	İ			•	VAPILSMFFLMCYLFVNLACALQTLLRTPNW
			[			RPRFRYYHWALSFMGMSICLALMFISSWYYA
						IVAMVIAGMIYKYIEYQGAEKEWGDGIRGLS
		l			İ	LSAARFALLRLEEGPPHTKNWRPQLLVLLKL
1	1 .	1			J	DEDLHVKHPRLLTFASQLKAGKGLTIVGSVIV GNFLENYGEALAAEQTIKHLMEAEKVKGFCQ
1	ļ .					LVVAAKLREGISHLIQSCGLGGMKHNTVVM
						GWPNGWRQSEDARAWKTFIGTVRVTTAAHL
						ALLVAKNISFFPSNVEQFSEGNIDVWWIVHDG
1						GMLMLLPFLLK\QHKVWRKCSIRFF\TVAOLE
	Į.					DNSIQMKKDLATFLYHLRIEAEVEVVEMHDS
1 .	j					DISAYTYERTLMMEQRSQMLRHMRLSKTER
1						DREAQLVKDRNSMLRLTSIGSDEDEETETYO
						EKVHMTWTKDKYMASRGQKAKSMEGFQDL
1						LNMRPDQSNVRRMHTAVKLNEVIVNKSHEA
						KLVLLNMPGPPRNPEGDENYMEFLEVLTEGL
812	2162	A	6600	-	610	ERVLLVRGGGSEVITTYS
012	2102	A	6628	66	640	AVCTMSEMAELSELYEESSDLQMDVMPGEG
1						DLPQMEVGSGSRELSLRPSRSGAQQLEEEGP
						MEEEEAQPMAAPEGKRSLANGPNAGEQPGQ VAGADFESEDEGEEFDDWEDDYDYPEEEQLS
						GAGYRVSAALEEADKMFLRTREPALDGGFQ
					-	MHYEKTPFDQLAFIEELFSLMVVNRLTEELG
						CDEIIDRE
813	2163	A	6630	708	1355	AKMGAYKYIQELWRKKQSDVMRFLLRVRC
						WQYRQLSALHRAPRPTRPDKARRLGYKAKQ
						GY/VYIYIGFVFAVIYRIRVRRGGRKRPVPKG
. 1		İ				ATYGKPVHHGVNQLKFARSLQSVAEERAGR
]		J				HCGALRVLNSYWVGEDSTYKFFEVILIDPFHK
			Ĭ	1		AIRRNPDTQWITKPVHKHREMRGLTSAGRKS
						RGLGKGHKFHHTIGGSRRAAWRRRNTLQLH RYR
814	2164	A	6635	201	1705	KGTEMNKSRWQSRRRHGRRSHQQNPWFRLR
1		*	0035	201	1,05	DSEDRSDSRAAQPAHDSGHGDDESPSTSSGT
1 1						AGTSSVPELPGFYFDPEKKRYFRLLPGHNNCN
] ]			1		}	PLTKESIRQKEMESKRLRLLQEEDRRKKIARM
			1			GFNASSMLRKSQLGFLNVTNYCHLAHELRLS
[ ]			1			CMERKKVOIRSMDPSALASDRFNLILADTNS
[ {	1	l	ı	ľ		DRLFTVNDVTVGGSKYGIINLQSLKTPTLKVF
		İ				MHENLYFTNRKV\NSVCWASLNHLDSHILLC
ļ [						LMGLAETPGCATLLPASLFVNSHPAGIDRPG\
j l	ļ				ļ	MLCSFRIPGAWSCAWSLNIQANNCFSTGLSR
[	1	[		ĺ	ſ	RVLLTNVVTGHRQSFGTNSDVLAQQFALMA
]	ļ		- 2			PLLFNGCRSGEIFAIDLRCGNQGKGWKATRLF
'	ł	ł	ł	- 1	ł	HDSAVTSVRILQDEQYLMASDMAGKIKLWD LRTTKCVRQYEGHVNEYAYLPLHVHEEEGIL
]	l		. [		l	VAVGQDCYTRIWSLHDARLLRTIPSPYPASKA
ļ	l		[			DIPSVAFSSRLGGSRGAPGLLMAVGQDLYCY
			• !			SYS
815	2165	A	6643	659	3282	NKNILEVPSARTTRIMGDHLDLLLGVVLMAG
·	i		j	1		PVFGIPSCSFDGRIAFYRFCNLTQVPQVLNTTE
		ł	į	ţ	j	RLLLSFNYIRTVTASSFPFLEQLQLLELGSOYT
]	-	1	{	l		PLTIDKEAFRNLPNLRILDLGSSKIYFLHPDAF
] }		ţ			ļ	QGLFHLFELRLYFCGLSDAVLKDGYFRNLKA
		1		J	İ	LTRLDLSKNQIRSLYLHPSFGKLNSLKSIDFSS
	Ì	ľ	ľ	ì	l	NQIFLVCEHELEPLQGKTLSFFSLAANSLYSR
	ļ			l	ĺ	VSVDWGKCMNPFRNMVLEILDVSGNGWTV
	}	I		i		DITGNFSNAISKSQAFSLILAHHIMGAGFGFHN
						IKDPDQNTFAGLARSSVRHLDLSHGFVFSLNS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  RVFETLKDLKVLNLAYNKINKIADEAFYGLD NLQVLNLSYNLLGBLYSSNFYGLPKVAYIDL QKNHIAIIQDQTFKFLEKLQTLDLRDNALTTIH FIPSIPDIFLSGNKLVTLPKINLTANLIHLSENR LENLDILYFLLRVPHLQILILNQNRFSSCSGDQ TPSENPSLEQLFLGENMLQLAWETELCWDVF EGLSHLQVLYLNHNYLNSLPPGVFSHLTALR GLSLNSNRLTVLSHNDLPANLEILDISRNQLL APNPDVFVSLSVLDITHNKFICECELSTFINWL NHTNVTIAGPPADIYCVYPDSLSGVSLFSLSTE GCDEEEVLKSLKFSLFIVCTVTLTLFLMTILTV TKFRGFCFICYKTAQRLVFKDHPQGTEPDMY KYDAYLCFSSKDFTWQNALLKHLDTQYSD QNRFNLCFEERDFVPGENRPANIQDAIWNSR KIVCLVSRHFLRDGWCLEAFSYAQGRCLSDL NSALIMVVVGSLSQYQLMKHQSIRGFVQKQQ YLRWPEDLQDVGWFLHKLSQOILKKEKEKK
816	2166	A	6646	1	3811	KDNNIPLQTVATIS  RDRAGVRPAGKQHAAAAFYDVGGDRPWDS  GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS
816	2166	A	6646		3811	RDRAGVRPAGKQHAAAAFYDVGGDRPWDS GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAAGHGNEFFQYTAPKQP KKGQGTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQQPVTVARIHVNFELMVRPNNY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEBLDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ
					-	EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDWLG
817	2167	<u> </u>	6649	63	1073	FFRSSSDNGSPIRQYE/HSTPAHQGPVMGLEG

SEQ ID NO; of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine
seq-	uence	ľ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Ghitamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		j .	j	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ			peptide		/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
						KS/ARNSQLRIVLVGKTGAGKSATGNSILGRK
1	ļ			Į.		VFHSGTAAKSITKKCEKRSSSWKETELVVVD
				ļ.	1	TPGIFDTEVPNAETSKEURCILLTSPGPHALLL
					İ	VVPLGRYTEEEHKATEKILKMFGERARSFMIL
1						IFTRKDDLGDTNLHDYLREAPEDIQDLMDIFG
[						DRYCALNNKATGAEQEAQRAQLLGLIQRVV
					ĺ	RENKEGCYTNRMYQRAEEEIQKQTQAMQEL
						HRVELEREKARIREEYEEKIRKLEDKVEQEKR
						KKQMEKKLAEQEAHYAVRQQRARTEVESKD
				·		GILELIMTALQIASFILLRLFAED
818	2168	A	6660	357	1890	APSGSWTRVVLTLDPCSLRSRSPRSLLDPGMP
						GISARGLSHEGRKQLAVNLTRVLALYRSILDA
1 1					0.1	YHEFF\TDNLWDTLPCSWQEALDGLKPPQLA
٠٠.	,					TMLLGMPGEGEVVRYRSVWPLTLLALKSTA
1 7						CALAFTRMPGFQTPSEFLENPSQSSRLTAPFR
1						KHVRPKKQHEIRRLGELVKKLSDFT/GLHPGC RRGLRPGVHLSRFMALGLGLMVKSIEGDQRL
						VERAQRLDQELLQALEKEEKRNPQVVQTSPR
j j						HSPHHVVRWVDPTALCEELLLPLENPCOGRA
					•	RLLLTGLHACG\DLSVALLRHFSCCPEVVALA
						SVGCCYMKLSDPGGYPLSQWVAGLPGYELP
						YRLREGACHALEEYAERLQKAGPGLRTHCY
i i		· 1			i	RAALETVIRRARPELRRPGVQGIPRVHELKIEE
		1				YVQRGLQRVGLDPQLPLNLAALQAHLAQEN
		1				RVVAFFSLALLLAPLVETLILLDRLLYLQEQA
	J	J				LSP\GFHAELLPIFSPELSPRNLVLVATKMPLG
100	01/0					QALSVLETEDS
819	2169	A	6661	65	2686	SGSGHCLAEAASMGPWGWKLRWTVALLLA
7	•	1				AAGTAVGDRCERNEFQCQDGKCISYKWVCD
	- 1	ı		ł		GSAECQDGSDESQETCLSVTCKSGDFSCGGR
						VNRCIPQFWRCDGQVDCDNGSDEQGCPPKTC SQDEFRCHDGKCISRQFVCDSDRDCLDGSDE
		1	- 1			ASCPVLTCGPASFQCNSSTCIPQLWACDNDPD
]		i				CEDGSDEWPQRCRGLYVFQGDSSPCSAFEFH
			j		·	CLSGECIHSSWRCDGGPDCKDKSDEENCAVA
		I	]			TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS
		İ	į	ŀ	ļ	DEVGCVNVTLCEGPNKFKCHSGECITLDKVC
	.	Į	l		}	NMARDCRDWSDEPIKECGTNECLDNNGGCS
		į	1		1	HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE
		l			1	CODPDTCSQLCVNLEGGYKCQCEEGFQLDPH
		1				TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY
ľ	ſ	i	ĺ	į	1	TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI
	- 1	į		. [	İ	CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD
			1			WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR
	J	ļ	. ]	j	J	ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK
	.	·	i		1	KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL
	ł	į				YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH
	ŀ			į		PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV
	ł	l	- 1	1	ł	NLLAENLLSPEDMVLFHNLTQPRGVNWCERT
	1		į	l	1	TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM
.	ĺ	l	1		ļ	LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS
	j	ļ	J	J	J	STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI
[	1	l	J	ſ	ſ	VTMSHQALGDVAG\RGN\EKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP
1	į	l		l		VYQKTTEDEVHICHNQDGYSYPSRQMVSLED
	l	l	l			DVA DVA
820	2170	A	6666	17	4146	ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA
1		-			.270	LSLWPTSGEICGPGIDIRNDYQQLKRLENCTVI
			1		l	EGYLHILLISKAEDYRSYRFPKLTVITEYLLLF
	l	i	l			RVAGLESLGDLFPNLTVIRGWKLFYNYALVIF

1					<u>.</u>	VGGQNPSTGGISADRTQGNIGCGGDTDPGQS
						PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS
823	2173	A	6727	3		SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM
						TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC
		1	. [	}		QISKLQKEFKR\HINSDAHSTTS\SASP\AOSPLF
	İ	.				VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST
		- 1	- 1	ļ		SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH
624	2172	A	6715	772	21	DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKRYEK\DFQ\SCFGLHETR\
822	2172		6715	777		RREERLQAKKEEIIKTLSKEEETKK
						HKLKADKARKKLLADQAEARRSKTKEARK
	1		- 1	Ì	ŀ	ARMPEKVTWMRRMRILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH
						VHSRARCRKNTLARRKGRHMGIGKRKGTAN
		ľ	- 1	,	ł	SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT
821	2171	^	6691	106	825	GRVLFRGCGVGHKGQVLMGTFILAQDWLSE
921	2171		((0)	106		DERQPYAHMNGGRKNERALPLPQSSTC
	1					VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF
	1	1	ſ		1	MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE
		Į		[	. 1	GLSNEQVLRFV\MEGGLLDKPDNCPDMLFEL
					ļ	FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ
			{		· [	AYLNANKFVHRDLAARNCMVAEDFTVKIGD
		J	J	j	ļ	SLRPEMENNPVLAPPSLSKMIQMAGEIADGM
			ŀ		[	AIKTVNEAASMRERIEFLNEASVMKEFNCHH VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR
] ]			j			MSRELGQGSFGMVYEGVAKGVVKDEPETRV
] :						GVLYASVNPEYFSAADVYVPDEWEVAREKIT
			ļ		ļ	QATSLSGNGSWTDPVFFYVQAKRYENFIHLII ALPVAVLLIVGGLVIMLYVFHRKRNNSRLGN
			l			DQRECVSRQEYRKYGGAKLNRLNPGNYTARI
					}	RPENSIFLKWPEPENPNGLILMYEIKYGSOVE
						ESRVDNKERTVISNLRPFTLYRIDIHSCNHEAE KLGCSASNFVFARTMPAEGADDIPGPVTWEP
1	]		1	ļ		NTTMSSRSRNTTAADTYNITDPEELETEYPFF
					•	AEYRKVFENFLHNSIFVPRPERKRRDVMQVA
	[					QDGYLYRHNYCSKDKIPIRKYADGTIDIEEVT ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE
1				į		NSSSQLIVKWNPPSLPNGNLSYYIVRWQROP
1	j					VENDHIRGAKSEILYIRTNASVPSIPLDVLSAS
1						EAPFKNVTEYDGQDACGSNSWNMVDVDLPP NKDVEPGILLHGLKPWTQYAVYVKAVTLTM
	] [	]				TSTTTSKNRIIITWHRYRPPDYRDLISFTVYYK
1						VTGTKGRQSKGDINTRNNGERASCESDVLHF
1	]			,		NLRLILGEEQLEGNYSFYVLDNQNLQQLWD WDHRNLTIKAGKMYFAFNPKLCVSEIYRMEE
						SELENFMGLIEVVTGYVKIRHSHALVSLSFLK
	1					KTIDSVTSAQMLQGCTIFKGNLLINIRRGNNIA
						RCVDRDFCANILSAESSDSEGFVIHDGECMQE CPSGFIRNGSQSMYCIPCEGPCPKVCEEEKKT
		}				DTACVACRHYYYAGVCVPACPPNTYRFEGW
						MCPSTCGKRACTENNECCHPECLGSCSAPDN
						STVDWSLILDAVSNNYIVGNKPPKECGDLCP GTMEEKPMCEKTTINNEYNYRCWTTNRCQK
					]	EMTNLKDIGLYNLRNITRG\AIRIEKNADLCYL
<b> </b>	<u> </u>		<b> </b>	sequence		nucleotide insertion
1				peptide	- Sequence	/=possible nucleotide deletion, \=possible
				amino acid residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
nucl- eotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ĺ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ì	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
	1	ĺ		sequence		nucleotide insertion
			<del>                                     </del>	<del></del>		SSQPSQDGQESNVPSVGSLADPDYLNTPQMN
ı	1				1	TPVTLNSAAPASNSGAGVLPSPATPRFSVPTP
	l				l	RTPRTPRTPRGGGTASGQGSVKYDSTDQGSP
	l	l	1	ļ	<b>!</b>	ASTPSTTRPLNSVEPATMOPIPEAHSLYVTLIL
			ĺ			SDSVMNIFKDRNFDSCCICACNMNIKGADVG
	ļ			ŀ	1	
		İ	1			LYIPDSSNEDQYRCTCGFSAIMNRKLGYNSGL FLEDELDIFGKNSDIGQAAERRLM/MCQSTFL
			Ī			POTECTER DOEDDED I I I ONOREODE CENT
	{	ł		ł		PQVEGTKKPQEPPISLLLLLQNQHTQPFASLN
ļ	Į	1	1			FLDYISSNNRQTLPCVSWSYDRVQADNNDY
		1	i			WTECFNALEQGRQYVDNPTGGKVDEALVRS
		i			i .	ATVHSWPHSNVLDISMLSSQDVVRMLLSLQP
				l		FLQDAIQKKRTGRTWENIQHVQGPLTWQQFH
	[	ľ	1	f		KMAGRGTYGSEESPEPLPIPTLLVGYDKDFLT
	,	Ī	1			ISPFSLPFWERLLLDPYGGHRDVAYIVVCPEN
		l	1	ĺ		EALLEGAKTFFRDLSAVYEMCRLGQHKPICK
- 1			l	ľ		VLRDGIMRVGKTVAQKLTDELVSEWFNQPW
				·		SGEENDNHSRLKLYAQVCRHHLAPYLATLQL
1						DSSLLIPPKYQTPPAAAQGQATPGNAGPLAPN
			i			GSAAPPAGSAFNPTSNSSSTNPAASSSASGSSV
						PPVSSSASAPGISQISTTSSSGFSGSVGGQNPST
[						GGISADRTQGNIGCGGDTDPGQSSSQPSQDG
1						QESVTERERIGIPTEPDSADSHAHPPAVVTYM
- 0						VDPFTYAAEEDSTSGNFWLLSLMRCYTEMLD
						NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY
						IQYLKSMAFSVYCQCRRPLPTQIHIKSLTGFGP
						AASIEMTLKNPERPSPIQLYSPPFILAPIKDKQT
•			]			ELGETFGEASQKYNVLFVGYCLSHDQRWLL
			1			ASCTDLHGELLETCVVNIALPNRSRRSKVSAR
			1			KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR
l			}			LGHGELKDWSILLGECSLQTISKKLKDVCRM
			!			CGISAADSPSILSACLVAMEPQGSFVVMPDAV
			i I			TMGSVFGRSTALNMQSSQLNTPQDASCTHIL
J						VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL
						PFPDDMDNDIGILMTGNLHSSPNSSPVPSPGSP
						SGIGVGSHFQHSRSQGERLLSREAPEELKQQP
l					1	LALGYFVSTAKAENLPQWFWSSCPQAQNQC
į.						PLFLKASLHHHISVAQTDELLPARNSQRVPHP
ľ			1			LDSKTTSDVLRFVLEQYNALSWLTCNPATQD
	i					RTSCLPVHFVVLTQLYNAIMNIL
004						THE COLUMN TERMS
824	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR
024	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR
024	2174	A	6732	2440	365	
024	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP
024	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC
044	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD
024	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV
024	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG
<b>6</b> 24	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM
024	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGTTKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP
024	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGTTKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPTHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLRAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN
024	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK
624	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS
624	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL
624	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEBTTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI
624	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSFFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLISILIS SGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE
024	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGTTKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDDKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGAPPDLYCNEDSWQLPHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ
624	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGTTKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLSILIS SGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK
624	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGTTKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLSILIS SGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA
0.24	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYG\KLESL\SILIS SGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAYFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT
0.24	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGTTKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLSILIS SGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide delection, \=possible nucleotide insertion  QLPLPRSLHNYLLYEDVLRMYEVPELAAIQD G  RIMGLFDRGVQMLLTTVGAFAAFSLMTIAVG TDYWLYSRGVCKTKSVSENETSKKNEEVMT HSGLWRTCCLEGNFKGLCKQIDHFPEDADYE ADTAEYFLRAVRASSIFPILSVILLFMGGLCIA ASEFYKTRHNIILSAGIFFVSAGLSNIIGIIVYIS ANAGDPSKSDSKKNSYSYGWSFYFGALSFIIA EMVGVLAVHMFIDRHKQLRATARA\TDYLQ ASAITRIPSYRYRYQRRSRSSSRSTEPSHSRDA SPVGIKGFNTLPSTEISMYTLSRDPLKAATTPT ATYNSDRDNSFLQVHNCIQKENKDSLHSNTA
826	2176	A	6744	3		NRRTTPV  SDDLRTGLFQDVQDAESLKLPGVYEVLFYNE TEDCPGMMLWRYPEPRGLTLVRITPVFPNTT EDPDISTADLGDVLQDPCSLEYWDELQKVFV AFREFNLSESK VCELQLPDINLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSQSTCDPL VTPTALAACTRVDSCFTPWFVPSLCVSFQFAH LEFHLCHHLDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIQFL AQADCKLLECRNVTMQSVVKPFSIFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AIQAWQQNKCPEVEELVFSHFVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEFNS QVQQRMIVFSPLFIMRSHLPDPIIHLEKRSLGL SETQIIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQLID EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFBGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFTTVAPEGRAGPILTNINRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASBATTSISTPKERGPIFTTARQLVHALAMHY AAGALFRAGWVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSSPVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI AGIVDQPMQNFQKTSEAQASAGHKAKGVISG

SEQ ID NO: of nucl-ordide cotide sequence (Malanine Carty Carlotted and nucleotide location of peptide sequence (Malanine Carlotted and nucleotide location (Malanine Carlotted and nucleotide location (Malanine Carlotted and nucleotide location (Malanine Carlotted and nucleotide location (Malanine Carlotted and nucleotide location (Malanine Carlotted and nucleotide location (Malanine Carlotted and nucleotide location (Malanine Carlotted and nucleotide location (Malanine Carlotted and nucleotide location (Malanine Carlotted and nucleotide location (Malanine Carlotted and nucleotide location (Malanine Carlotted and nucleotide location (Malanine Carlotted and nucleotide location (Malanine Carlotte	ie,
nucleotide sequence   District	•
eotide sequence  USSN 09/496 914  USSN 09/496 914  Esoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proling additionated fresidue of peptide sequence  Frince sequence  USSN 09/496 914  Esoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proling additionated fresidue of peptide sequence  Frince sequence  USSN 09/496 914  Esoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proling additionated fresidue of peptide sequence  Fril remine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon /=possible nucleotide deletion, \=possible nucleotide insertion  VGKGMGVFTKPJGGAAELVSQTGY, GLSQLPKQRHQPSDVHADQAPNSHY KMLQSLGRPEVHMALDVVLVRGSG LLLTSEVLFVSVSEDTQQAFPVTE SKQNNLLTVQLKQPRVACDVEVDGY QQYNRLVDYTTKTSCHLAPSCSSMQI AEPPFSTVKTYHYLVDPHFAQVFLSK NKALRKGFP  FVGAPRRGNPFGSFGNPGRHQGPCH ASGVSPTLWRPQAAATGLEMPSSGR LDSGSLTSLDSSVFCSEGEGEPLALGI VGGSRFVLSQQALSCFPHTRLGKLAN RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAL QYWGIDELSIDSCCRDRYFRRKELSE DTEDQESQHESEQDFSQGPCFTVRQK EKPGSSTAARIFGVISIIFVGVSIINMA SWLDLQLLEILEYCISWFTGEFVLR RCRFLRKVPNIIDLLAILPFYITLLVES TQELENVGAHCPGCLRLIRALRMLHSTGLRSLGMTTTQCYEEVGLLLIFL STVEYFAEQSIPDTTFTSVPCAWWW TVGYGDRPDTTTGKIVAFMCILSGII AIINDRFSACYFTLKLKEAAVRQREA	•
sequence uence	E,
1	*
amino acid residue of peptide residue of peptide sequence peptide sequence peptide sequence s	
residue of peptide sequence  Y=Tyrosine, X=Unknown, *=Stop codon /=possible nucleotide deletion, \=possible nucleotide insertion  VGKGIMGVFTKPIGGAAELVSQTGYGGLSQLPKQRHQPSDVHADQAPNSHYKMLQSLGRPEVHMALDVVLVRGSGLLLLTSEVLFVVSVSEDTQQAFFVTE SKQNNLLTVQLKQPRVACDVEVDGYQAYPNLVDYITKTSCHLAPSCSSMQIQAYNRLVDYITKTSCHLAPSCSSMQIQAEPPSTVKTYHYLVDPHFAQVFLSKNKALRKGFP  827 2177 A 6748 2 1662 FVGAPRRGNPFGSPGNPGRHQGPCHASGVSPTLWRPQAAATGLEMPSSGRLDSGSLTSLDSSVFCSEGEGEPLALGIVGGSRFVLSQQALSCFPHTRLGKLANRPGALAAVPSPLELCDDANPVDNESQAFRYVLHYYRTGRLHVMEQLCALQYWGIDELSIDSCCRDRYFRRKELSEDTEDQESQHESEQDFSQGPCPTVRQKEKPGSSTAARIFGVISIIFVGVSIINMASWLDLLEILEYVCISWPTGEFVLRRCRFLRKVPNIIDLLAILFFYTTLLVESTQELVENVGAHCPGCLRLLRALNRMLHSTGLRSLGMTTTQCYEEVGLLLLFLSTVEYFAEQSIPDTTFTSVPCAWWWTVGYGDIRPDTTTGKVVAFMCILSGILAIINDRFSACYFTLKLKEAAVRQREA	
Peptide   Sequence	
827 2177 A 6748 2 1662 FVGAPRRGNPFGSPGNPGRHQGPCHI ASQUERYVLHYVRTGRIALGE BERGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAI QYWIDELSIDSCCRDRYFRRKELSE DTEDQESQHESQDPSQGPCTVRQV EKPGSTAARIFGVISIIFVGVSIINMA SWLDLQLLEILEYVCISWPTGEFVLRI RCRFLRKVPNIIDLLAILPFYITLLVES TQELWINGAHCPGCLRLLRALWML HSTGLRSLGMTTTGCYEEVGLLLIFL STVEYFAEQSIPDTTIFTSVPCAWWW TVGYGDIRPDTTIGKIVAFMCILSGIL AINDRFSACYFTLKLKEAAVRQREA	
VGKGIMGVFTKPIGGAAELVSQTGYY GLSQLPKQRHQPSD\VHADQAPNSHY KMLQSLGRPEVHMALDVVLVRGSG LLLTSEVLFVVSVSEDTQQQAFPVTE SKQNNLLTVQLKQPRVACDVEVDGY QQYNRLVDYTIKTSCHLAPSCSSMOJ AEPPPSTVKTYHYLVDPHFAQVFLSK NKALRKGFP  827 2177 A 6748 2 1662 FVGAPRRGNPFGSPGNPGRHQGPCH ASGVSPTLWRPQAAATGLEMPSSGR LDSGSLTSLDSSVFCSEGGEPLALGI VGGSRFVLSQQALSCFPHTRLGKLAY RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAI QYWGIDELSIDSCCRDRYFRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISIIFVGVSIINMA SWLDLQLLEILEYVCISWPTGEFVLR RCRFLRKVPNIIDLLAILPFYITLLVES TQELÆNVGAHCPGCLRLLRALRML HSTGLRSLGMTTTQCYEEVGLLLIFL STVEYFAEQSIPDTTFTSVPCAWWW TVGYGDIRPDTTTGKTVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
GLSQLPKQRHQPSD\VHADQAPNSHY KMLQSLGRPEVHMALDVVLVRGSG LLLTSEVLFVVSVSEDTQQQAFPVTE SKQNNLLTVQLKQPRVACDVEVDGY QQYNRLVDYTTKTSCHLAPSCSSMQI AEPPPSTVKTYHYLVDPHFAQVFLSK NKALRKGFP  827 2177 A 6748 2 1662 FVGAPRRGNPFGSPGNPGRHQGPCH ASGVSPTLWRPQAAATGLEMPSSGR LDSGSLTSLDSSVFCSEGEGEPLALGI VGGSRFVLSQQALSCFPHTRLGKLAV RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAL QYWGIDELSIDSCCRDRYFRRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISHFVGVSHNMA SWLDLQLLEILEYVCISWFTGEFVLRJ RCRFLRKVPNHIDLLAILPFYTILLVES TQELVENVGAHCPGCLRLLRALRMLJ HSTGLRSLGMTTTQCYEEVGLLLIFL STVEYFAEQSIPDTTTTSVPCAWWW TVGYGDRPDTTTGKIVAFMCILSGIL AHNDRFSACYFTLKLKEAAVRQREA	THE A
KMLQSLGRPEVHMALDVVLVRGSG- LLLTSEVLFVVSVSEDTQQAFPVTE SKQNNLLTVQLKQPRVACDVEVDGY QQYNRLVDYTTKTSCHLAPSCSSMQI AEPPPSTVKTYHYLVDPHFAQVFLSK NKALRKGFP  827 2177 A 6748 2 1662 FVGAPRRGNPFGSPGNPGRHQGPCH ASGVSPTLWRPQAAATGLEMPSSGR LDSGSLTSLDSSVFCSEGEGEPLALGI VGGSRFVLSQQALSCFPHTRLGKLAY RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAL QYWGIDELSIDSCCRDRYFRRKELSE DTEDQESQHESEQDPSQGPCPTVRQK EKPGSSTAARIFGVISIIFVGVSIINMA SWLDLQLLEILEYVCISWFTGEFVLRI RCRFLRKVPNIIDLLAILPFYTTLLVES TQELVENVGAHCPGCLRLLRALRRML HSTGLRSLGMTTTQCYEEVGLLLLFL STVEYFAEQSIPDTTTTSVPCAWWW TVGYGDIRPDTTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
LLLTSEVLFVVSVSEDTQQQAFPVTE SKQNNLLTVQLKQPRVACDVEVDGY QQYNRLVDYITKTSCHLAPSCSSMQI AEPPSTVKTYHYLVDPHFAQVFLSK NKALRKGFP  827 2177 A 6748 2 1662 FVGAPRRGNPFGSPGNPGRHQGPCHI ASGVSPTLWRPQAAATGLEMPSSGR LDSGSLTSLDSSVFCSEGEGEPLALGI VGGSRFVLSQQALSCFPHTRLGKLAV RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAI QYWGIDELSIDSCCRDRYFRRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISHFVGVSHNMA SWLDLQLLEILEYVCISWPTGEFVLRI RCRFLRKVPNHIDLLAHFFYTILLVES TQELVENVGAHCPGCLRLLRALRMLH HSTGLRSLGMTTTQCYEEVGLLLLFL STVEYFAEQSIPDTTFTSVPCAWWW TVGYGDIRPDTTTGKIVAFMCILSGIL AHNDRFSACYFTLKLKEAAVRQREA	
SKQNNLLTVQLKQPRVACDVEVDGY QQYNRLVDYITKTSCHLAFSCSSMQI AEPPPSTVKTYHYLVDPHFAQVFLSK NKALRKGFP  827 2177 A 6748 2 1662 FVGAPRGNPFGSPGNPGRHQGPCHI ASGVSPTLWRPQAAATGLEMPSSGR LDSGSLTSLDSSVFCSEGEGEPLALGI VGGSRFVLSQQALSCPPHTRLGKLAY RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAI QYWGIDELSIDSCCRDRYFRRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISHFVGVSHNMA SWLDLQLLEILEYVCISWPTGEFVLRI RCRFLRKVPNHIDLLALIFFYTILLVES TQELVENVGAHCPGCLRLLRALNRI HSTGLRSLGMTTTQCYEEVGLLLIFL STVEYFAEQSIPDTTFTSVPCAWWW TVGYGDIRPDTTTGKIVAFMCILSGIL AINDRFSACYFTLKLKEAAVRQREA	
QQYNRLVDYITKTSCHLAPSCSSMQI AEPPPSTVKTYHYLVDPHFAQVFLSK NKALRKGFP  827 2177 A 6748 2 1662 FVGAPRGNPFGSPGNPGRHQGPCH ASGVSPTLWRPQAAATGLEMPSSGR LDSGSLTSLDSSVFCSEGEGEPLALGI VGGSRFVLSQQALSCPPHTRLGKLAV RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAI QYWGIDELSIDSCCRDRYFRRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISIIFVGVSIINMAI SWLDLQLLEILEYVCISWFTGEFVLRI RCRFLRKVPNIIDLLAILPFYTILLVES TQELVENVGAHCPGCLRLRALVRMLI HSTGLRSLGMTTTQCYEEVGLLLFLL STVEYFAEQSIPDTTFISVPCAWWW TVGYGDIRPDTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
AEPPPSTVKTYHYLVDPHFAQVFLSK NKALRKGFP  827 2177 A 6748 2 1662 FVGAPRRGNPFGSPGNPGRHQGPCH ASGVSPTLWRPQAAATGLEMPSSGR LDSGSLTSLDSSVFCSEGEGEPLALGI VGGSRFVLSQALSCFPHTRLGKLAV RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAL QYWGIDELSIDSCCRDRYFRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISHFVGVSHNMA SWLDLQLLEILEYVCISWFTGEFVLR RCRFLRKVPNHIDLLAILFFYTILLVES TQELVENVGAHCPGCLRLLRALVRML HSTGLRSLGMTTTQCYEEVGLLLFL STVEYFAEQSIPDTTFTSVPCAWWW TVGYGDIRPDTTTGKIVAFMCILSGIL AHNDRFSACYFTLKLKEAAVRQREA	
NKALRKGFP  827 2177 A 6748 2 1662 FVGAPRRGNPFGSPGNPGRHQGPCH ASGVSPTLWRPQAAATGLEMPSSGR LDSGSLTSLDSSVFCSEGEGEPLALGI VGGSRFVLSQQALSCFPHTRLGKLAV RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAI QYWGIDELSIDSCCRDRYFRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISIIFVGVSIINMAI SWLDLQLLEILEYVCISWPTGEFVLRI RCRFLRKVPNIIDLLAILFFYITLLVES TQELVENVGAHCPGCLRLLALRMLI HSTGLRSLGMTTTQCYEEVGLLLFLI STVEYFAEQSIPDTTFTSVPCAWWW, TVGYGDRPDTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
827 2177 A 6748 2 1662 FVGAPRRGNPFGSPGNPGRHQGPCHI ASGVSPTLWRPQAAATGLEMPSSGR LDSGSLTSLDSSVFCSEGEGEPLALGI VGGSRFVLSQQALSCFPHTRLGKLAV RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAI QYWGIDELSIDSCCRDRYFRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISHFVGVSHNMA SWLDLQLLEILEYVCISWFTGEFVLRJ RCRFLRKVPNHIDLLAILPFYHTLLVES TQELVENVGAHCPGCLRLLRALKMLJ HSTGLRSLGMTTTQCYEEVGLLLFL STVEYFAEQSIPDTTFTSVPCAWWW TVGYGDRPDTTTGKIVAFMCILSGIL AHNDRFSACYFTLKLKEAAVRQREA	TIMAK
ASGVSPTLWRPQAAATGLEMPSSGR LDSGSLTSLDSSVFCSEGEGEPLALGI VGGSRFVLSQQALSCFPHTRLGKLAV RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAL QYWGIDELSIDSCCRDRYFRRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISIIFVGVSIINMA SWLDLQLLEILEYVCISWPTGEFVLRI RCRFLRKVPNIIDLLAILPFYITLLVES TQELVENVGAHCPGCLRLLRALVRMLL HSTGLRSLGMTTTQCYEEVGLLLLFL STVEYFAEQSIPDTTFTSVPCAWWW TVGYGDIRPDTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	DD COTT
LDSGSLTSLDSSVFCSEGEGEPLALGI VGGSRFVLSQQALSCFPHTRLGKLAV RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAI QYWGIDELSIDSCCRDRYFRRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISHFVGVSHMAA SWLDLQLLEILEYVCIS WPTGEFVLRI RCRFLRKVPNHDLLAILPFYTILLVES TQELWNVGAHCPGCLRLLRALRMLL HSTGLRSLGMTITQCYEEVGLLLLIFL STVEYFAEQSIPDTTFTSVPCAWWW TVGYGDIRPDTTTGKIVAFMCILSGIL AHNDRFSACYFTLKLKEAAVRQREA	
VGGSRFVLSQQALSCFPHTRLGKLAN RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAI QYWGIDELSIDSCCRDRYFRRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISHFVGVSHMAA SWLDLQLLEILEYVCISWPTGEFVLRI RCRFLRKVPNHDLLAILPFYTILLVES TQELWNVGAHCPGCLRLLRALRMLI HSTGLRSLGMTITQCYEEVGLLLLFL STVEYFAEQSIPDTTFTSVPCAWWW TVGYGDIRPDTTTGKIVAFMCILSGIL AHNDRFSACYFTLKLKEAAVRQREA	
RRPGALAAVPSPLELCDDANPVDNE SQAFRYVLHYYRTGRLHVMEQLCAI QYWGIDELSIDSCCRDRYFRRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISIIFVGVSIINMA SWLDLQLLEILEYVCISWPTGEFVLRI RCRFLRKVPNIIDLLAILPFYITLLVES TQELVENVGAHCPGCLRLLRALVRMLI HSTGLRSLGMTITQCYEEVGLLLLFLI STVEYFAEQSIPDTTFTSVPCAWWW TVGYGDIRPDTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
SQAFRYVLHYYRTGRLHVMEQLCAI QYWGIDELSIDSCCRDRYFRRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISHFVGVSINMA SWLDLQLLEILEYVCISWPTGEFVLRI RCRFLRKVPNHIDLLAILPFYTILLVES TQELVENVGAHCPGCLRLLRALVRMLI HSTGLRSLGMTTTQCYEEVGLLLLFL STVEYFAEQSIPDTTFTSVPCAWWW TVGYGDIRPDTTTGKIVAFMCILSGIL AHNDRFSACYFTLKLKEAAVRQREA	
QYWGIDELSIDSCCRDRYFRRKELSE DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISIIFVGVSIINMA SWLDLQLLEILEYVCISWPTGEFVLRI RCRFLRKVPNIIDLLAILPFYTILLVES TQEL\ENVGAHCPGCLRLLRAL\RMLI HSTGLRSLGMTTTQCYEEVGLLLLFL STVEYFAEQSIPDTTFTSVPCAWWW TVGYGDIRPDTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
DTEDQESQHESEQDFSQGPCPTVRQK EKPGSSTAARIFGVISIIFVGVSIINMAI SWLDLQLLEILEYVCISWPTGEFVLRI RCRFLRKVPNIIDLLAILFFYITLLVES TQELVENVGAHCPGCLRLLRALVRMLI HSTGLRSLGMTTTQCYEEVGLLLFLI STVEYFAEQSIPDTTFTSVPCAWWW, TVGYGDIRPDTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
EKPGSSTAARIFGVISIIFVGVSIINMA SWLDLQLLEILEYVCISWFTGEFVLRJ RCRFLRKVPNIIDLLAILPFYITLLVES TQELVENVGAHCPGCLRLLRALLX HSTGLRSLGMTITQCYEEVGLLLFL STVBYFAEQSIPDTTFTSVPCAWWW, TVGYGDIRPDTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
SWLDLQLLEILEYVCISWFTGEFVLRJ RCRFLRKVPNIIDLLAILPFYITLLVES TQELVENVGAHCPGCLRLLRAL\RMLJ HSTGLRSLGMTTTQCYEEVGLLLLFL STVEYFAEQSIPDTTFTSVPCAWWWA TVGYGDIRPDTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
RCRFLRKVPNIIDLLAILPFYTTLLVES TQELVENVGAHCPGCLRLLRALVRML; HSTGLRSLGMTTTQCYEEVGLLLLFL; STVEYFAEQSIPDTTFTSVPCAWWW, TVGYGDIRPDTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
TQEL\ENVGAHCPGCLRLLRAL\RML; HSTGLRSLGMTTTQCYEEVGLLLLFL; STVEYFAEQSIPDTTFTSVPCAWWW, TVGYGDIRPDTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
HSTGLRSLGMTTTQCYEEVGLLLLFL: STVEYFAEQSIPDTTFTSVPCAWWW/ TVGYGDIRPDTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
STVEYFAEQSIPDTTFTSVPCAWWW/ TVGYGDIRPDTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
TVGYGDIRPDTTTGKIVAFMCILSGIL AIINDRFSACYFTLKLKEAAVRQREA	
AIINDRFSACYFTLKLKEAAVRQREA	
I I I I I I I I I I I I I I I I I I I	
	KGKEK
The state of the s	
TASSGFQSMHSSNPKVRSSPSGNTQS:	
EVMVRPTVMSPSGNPQLDSKFSNQQ	
ASQSQPSPCDSKSGGHTPKALPGPGG	
NGAGNGAKGKRERSISADSFDQR	
DDSDIKECNSADHIKSQDSQHTPHSM	
APRSSTPPHGQTTATEPTPAQKTPAK	VVYVFS
TEMANKAABAVLKGQVETIVSFHIQN	ISNNK
TERSTAPLNTQISALRNDPKPLPQQPP	
DQNSSQNTRLQPTPPIPAPAPKPAAPP	
SPGVENKLIPSVGSPASSTPLPPDGTG	
NRAVTPVSQGSNSSSADPKAPPPPPVS	
LGENPDGLSQEQLEHRERSLQTLRDIG	
DEKEPTGAQSGGPQQNPGVLDGPQKI QAMMAQSQSLGKGPGPRTDVGAPFG	
DVPFSPDEMVPPSMNSQSGTIGPDHLI FOLANT VI OOFFVERVERVERVERVERVERVERVERVERVERVERVERVERV	
EQIAWLKLQQEFYEEKRRKPEQVVVQ	ACCOUNT
DMMVHQHGPRGVVRGPPPPYQMTPS	EUWAP
GGTEPFSDGINMPHSLPPRGMAPHPN	
MRLPGFAGMINSEMEGPNVPNPASRE	
SWPDDVPKIPDGRNFPPGQGIFSGPGR	
NPQGLSEEMFQQQLAEKQLGLPPGM	
PSMEMNRMIPGSQRHMEPGNNPIFPR	
LSPSRGDFPKGIPPQMGPGRELEFGMY	
KGDVNLNVNMGSNSQMIPQKMREAC	
MLKLRPGGSDMLPAQQKMVPLPFGE	
YGMGPRPFLPMSQGPGSNSGLRNLRE	
RTNSRLSHMPPLPLNPSSNPTSLNTAP	PIGPDQ
LGRKPLDISVAGSQVHSPGINPLKSPT	PVQRG
SPMLGSPSGNLKSPQTPSQLAGMLAG	PVQRG MIHQVQ
SIKSPPVLGSAAASPVHLKSPSLPAPSF	PVQRG MHQVQ PAAAA
PEPPLQSPGIPPNHKAPLTMASPAMLO	PVQRG MHQVQ PAAAA GWTSS
GPPPPTASQPASVNIPG\SLPSSTPYTM	PVQRG MHQVQ PAAAA GWTSS NVESG
SQNPLSIMMSR\MSKFAMPS\SNPGY	PVQRG MHQVQ PAAAA GWTSS NVESG PPEPTL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  KTVASSDDDSPPARSPNLPSMNNMPGMGINT QNPRISGPNPVVPMPTLSPMGMTQPLSHSNQ MPSPNAVGPNIPPHGVPMGPGLMSHNPIMGH GSQEPPMVPQGRMGFPQGFPPVQSPPQQVPFP HNGPSGGQGSFPGGMGFPGEGPLGRPSNLPQ SSADAALCKPGGPGGPDSFTVLGNSMPSVFT DPDLQEVIRPGATGIPEFDLSRIIPSEKPSQTLQ YFPRGEVPGRKQPQGPGPFSHMQGMMGEQ APRMGLALPGMGGPGPVGTPDIPLGTAPSMP GHNPMRPPAFLQQGMMGPHHRMMSPAQST
						MPGQPTLMSNPAAAVGMIPGKDRGPAGLYT
829	2179	A	6797	433	3	HPGPVGSPGMMMSMQGMMGP\NRTS
	~117	T.	0131		3	ASFFNFSICICKIILEVGPPVGHPAHDDVGGRH GPGGR/GSRSPRSLQCAPGGGRRSGCPAGSSP ASTCPPSPGGSGADRFGPSPPPPSREAAPTAG AAASSTSSGASCPPVPASSRWGVRSRTRSGSG GEREPRDRPSERPRLV
830	2180	A	6800	3	1911	LPERAFGPRTPRAPRRRRRILLLSPPPRPPPPL DREPRAPGPWLCPSRAGTAQDPARIRERRGR VAGGAAGPAMELRARGWWLLCAAAALVAC ARGDPASKSRSCGEVRQIYGAKGFSSS\DVPQ AEISGEHLRICPQGYTCCTSEMEENLANRSHA ELETALRDSSRVLQAMLATQLRSFDDHFQHL LNDSERTLQATFPGAFGELYTQNARAFRDLY SELRLYYRGANLHLEETLAEFWARLLERLFK QLHPQLLLPDDYLDCLGKQAEALRPF\GEAP\ RELRLAT\RA\FVAAR\SFVQGLGVAS\DVVR KVAQVPLG\PEC\SRAVIEAGSYC\ALHCVGVP GARPCPDYCRNVLKGCLANQADLDAEWRNL LDSMVLITDKFWGTSGVESVIGSVHTWLAEA INALQDNRDTLTAKVIQGCGNPKVNPQGPGP EEKRRGKLAPRERPPSGTLEKLVSEAKAQL RDVQDFWISLPGTLCSEKMALSTASDDRCWN GMARGRYLPEVMGDGLANQINNPEVEVDIT KPDMTIRQQIMQLKIMTNRLRSAYNGNDVDF QDASDDGSGSGSGDGCLDDLCGRKVSRKSSS SRTPLTHALPGLSEQEGQKTSAASCPQPPTFL LPLLLFLALTVARPRWR
831	2181	<b>A</b>	6808	2	1522	ASRHGMTPGALLMLLGALGPPLAPGVRGSEA EGRLREKLFSGYDSSVRPAREVGDRVRVSVG LILAQLISLNEKDEEMSTKVYLDLEWTDYRLS WDPAEHDGIDSLRITAESVWLPDVVLLNNND GNFDVALDISVVVSSDGSVRWQPPGIYRSSCS IQVTYFPFDWQNCTMVFSSYSYDSSEVSLQT GLGPDGQGHQEIHHEGTTEINGQWENHKPS RLIQPPGDPRGGREGQRQEVIFYLIIRRKPLFY LVNVIAPCILITLLAIFVFYLPPDAGEKMGLSIF ALLTLTVFLLLLADKVPETSLSVPIIIKYLMFT MVLVTFSVILSVVVLNLHHRSPHTHQMPLWV RQIFIHKLPLYLRLKRPKPERDLMPEPPHCSSP GSGWGRGTDEYFIRKPPSDFLFPKPNRFQPEL SAPDLRRFIDGPNRAVALLPELREVVSSISYIA RQLQEQEDHDALKEDWQFVAMVVDRLFLW TFIIFTSVGTLNVIFLDATYHLPPPDFFP
832	2182	A	6824	71	1079	ETMAKNPPENCEDCHILNAFAFKSKKICKSLK ICGLVFGILALTLIVLFWGSKHFWPEVPKKAY DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGN GTDETLEVHDFKNGYTGIYFVGLQKCFIKTQI KVIPEFSEPEEEIDENEETTTIFFEQSVIWVPAE KPIENRDFLKNSKILEICDNVTMYW\INPTL\IS

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	i	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
ucnce	1	1.	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	[			amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	Ì			peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		}		sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
			<del> </del>	Joquino		GTFAKQLHHNFAFIILVSELQDFEEEGEDLHFP
					i	ANEKKGIEQNEQWVVPQVKVEKTRHARQAS
ł		ļ		l	}	EEELPINDYTENGIEFDPMLDERGYCCIYCRR
						GNRYCRRVCEPLLGYYPYPYCYQGGRVICRV
						IMPCNWWVARMLGRV
833	2183	Α	6846	116	602	EAEGEQVCGAKCCGDAPHVENREEETARIGP
[ ]	'			ĺ		GVMESKEERALNNLIVENVNQENDEKDEKE
1						QVANKGEPLALPLNVSEYCVPRGNRRRFRVR
1					İ	QPILQYRWDIMHRLGEPQARMREENMERIGE
						EVRQLMEKLREKQLSHSLRAVSTDPPHHDHH
834	2184	Α	6851	3	2024	DEFCLMP
0.54	2104	Λ.	0021	3	2024	PNGVALLHLPGAAVIPNTNYMFQDALGGRSR
1 1		- 8-				GSREESPAPSRAPASASLWRRLVVVEAKMAA
]						HAAAAAQAAAAQAAHAEAADSWYLALLGF AEHFRTSSPPKIRLCVHCLQAVFPFKPPQRIEA
1 [						RTHLQLGSVLYHHTKNSEQARSHLEKAWLIS
ĺ						QQIPQFEDVKFEAASLLSELYCQENSVDAAKP
[ [						LLRKAIQISQQTPYWHCRLLFQLAQLHTLEKD
						LVSACDLLGVGAEYARVVGSEYTRALFLLSK
						GMLLLMERKLQEVHPLLTLCGQIVENWQGN
					•	PIQKESLRVFFLVLQVTHYLDAGQVKSVKPC
]						LKQLQQCIQTISTLHDDEILPSNPADLFHWLP
1						KEHMCVLVYLVTVMHSMQAGYLEKAQKYT
[ [		ĺ				DKALMQLEKLKMLDCSPILSSFQVILLEHIIM CRIVECTOR ATTALON CONTROL CON
					i	CRLVTGHKATALQEISQVCQLCQQSPRLFSN HAAQLHTLLGLYCVSVNCMDNAEAQPTTAL
	l l			1		RLTNHQELWAFIVTNLASVYIREGNRHQEVV
				Ì		LYSLLERINPDHSFPVSSHCLRAAAFYVRGLF
					·	SFFQGRYNEAKRFLRETLKMSNAEDLNRLTA
l						CSLVLLGHIFYVLGNHRESNNMVVPAMQLAS
ĺ		- 1	i			KIPDMSVQLWSSALLRDLNKACGNAMDAHE
		]				AAQMHQNFSQQLLQDHIEACSLPEHNLITWT
835	2185	A	6855	224	10/0	DGPPPVQFQAQNGPNTSLASLL
633	2105	^	0022	334	1268	PTRRPILPLTSPKAISVPSPLQGKQHTLVKSCL
						SVSGIGGFLVSLSSRMKLQTLAVSVTALKFWS
			ŀ	1		AYVPCQTQDRDALRLTLEQIDLIRRMCASYSE LELVTSAKALNDTQKLACLIGVEGGHSLDNS
			ſ	İ	ſ	LSILRTFYMLGVRYLTLTHTCNTPWAESSAK
				İ		GVHSFYNNISGLTDFGEKVVAEMNRLGMMV
İ	i	. !	l	ļ	. 1	DLSHVSDAVARRALEVSQAPVIFSHSAARGV
		ļ	- 1		ļ	CNSARNVPDDILQLLEEERWAFVMVSLFHGE
		1				LIQWQPIRPMCSTVADHFDHIKAV\IGSKFIGI
836	2106		(8/2	-315		GGDYDGAGKYRKKTTCKAPWRTSSRMSS
930	2186	A	6862	315	11	PPRSRPSCWRKKVGPGRPWWWGGTGPPGQG
		. 1		l	Ì	RPEIRLLPLPMTGACGAVAASRTGSSGPG/SSL
	ł		ł	ļ	1	PNGHGGKGSGLANGLAGNP\GHLGLGSSFGT GPGSGRPPP
837	2187	$\overline{\mathbf{A}}$	6863	2	1615	
	/	*		~	1013	VLRGQRGPAGGLAEERRGRNEWRIHDVIT APFPGLVQRRSRLLIVSQVRYFLKNKVSPDLC
		l		į		NEDGLTALHQCCIDNFEEIVKLLLSHGANVN
1	- 1		İ	t	ł	AKDNELWTPLHAAATCGHINLVKILVQYGA
Ì	l			i	. ]	DLLAVNSDGNMPYDLCEDEPTLDVIETCMAY
	1	- 1			ĺ	QGITQEKINEMRVAPEQQMIADIHCMIAAGO
			1	l	ł	DLDWIDAQGATLLHIAGANGYLRAAELLLDH
	j		- 1	ļ	1	GVRVDVKDWDGWEPLHAAAFWGQMQMAE
		1		l	į	LLVSHGANLNARTSMDEMPIDLCEEEEFKVL
	ł	- 1	- 1	ł		LLELK\HKHDVIMKSQLRHKSSLSRRTSHRQA
	j	1				S/SVGKVVRRTQPVGTGPNL\YRKEYE/GERAI
	<u>l</u>					LWQRSA\AEDQRTSTYNGDIRET\RTDQENKD

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PNPRLEK\PVLLSEFPTKIPRGELDMPVENGLR APVSAYQYALANGDVWKVHEVPDYSMAYG NPGVADATPPWSSYKEQSPQTILIELKRQRAA AKLLSHPFLSTHLGSSMARTGESSSEGKAPLI GGRTSPYSSNGTSVYYTVTSGDPPILIKFKAPI
838	2188	Α	6865	6291	739	EEMEEKVHGCCRIS  AGPLEPRVQGAMALQLWALTLLGLLGAGAS LRPRKLDFFRSEKELNHLAVDEASGVVYLGA VNALYQLDAKLQLEQQVATGPVLDNKKCTP PIEASQCHEAEMTDNVNQLLLVDPPRKRLVE
						CGQLLKGICALRALSMISLRLFYEDGSGEKSF VASNDEGVATVGLVSSTGPGGDRVLFVGKG NGPHDNGIIVSTRLLDRTDSREAFEAYTDHAT YKAGYLSTNTQQFVAAFEDGFYVFFVFNQQD KHPARNRTLLARMCREDPYVYSYLEMDLQC RDPDIHAAAFGTCLAASVAAPGSGRVLYAVF SRDSRSSGGPGAGLCLFFLDEVHAKMEANRN ACYTGTREARDIFYKPFHGDIQCGGHAPGSSK SFPCGSEHLPYPLGSRDGLRGTAVLQRGGLN LTAVTVAAENNHTVAFLGTSDGRILKVYLTP DGTSSEYDSILVEINKRVKRDLVLSGDLGSLY AMTQDKVFRLPVQECLSYPTCTQCRDSQDPY CGWCVVEGRCTRKAECPRAEEASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCLFGESPPHPARVEGEAVICNSPSS IPVTPPGQDHVAVTIQLLLRRGNIFLTSYQYPF YDCRQAMSLEENLPCISCVSNR WTCQWDLR YHECREASPNPEDGIVRAHMEDSCPQFLGFSP LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD LLKFMEPVTMQESGTFAFRTPKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGQSRCVYEALCNITT SECPPPVITIRIQPETGPLGGGRITTLGSNLGVQ AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP KPLSVEPQQGPQAGGTTLTIHGTHLDTGSQED VRVILNGVPCKVTKFGAQLQCVTGPPQATRG QMLLEVSYGGSPVPNPGIFTYRENPVLRAFE PLRSFASGGRSINVTQGGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDFTTENFTGGVKKQVNKLIRAR GTNLNKAMTLQBAEAFVGABRCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDCKKREFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPPERRPVVEQALYQFSNLLNSKSFLINFIHT LENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VRYAPLTVSVIVQDEGVBARVVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE RTKAITEIYLTRLLSVKGTLQQFVDNFFQSVL APGHAVPPAVKYFFDFLDEQAEKHNIQDEDTI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	i	USSN	location		
seq-	uence	[	09/496		corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	ucite			correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	ļ.	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	i	i .	ŀ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		1	· ·	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	i	1	ł	peptide		/=possible nucleotide deletion, \=possible
<u> </u>	L			sequence		nucleotide insertion
	[ <u></u>					HIWKTNSLPLRFWVNILKNPHFIFDVHVHEVV
	1		1	İ		DASLSVIAQTFMDACTRTEHKLSRDSPSNKLL
1	i	ĺ	i	i	1	YAKEISTYKKMVEDYYKGIRQMVQVSDQDM
ì		l				NTHLAEISRAHTDSLNTLVALHQLYQYTQKY
		Ì	]			
1			i			YDEIINALEEDPAAQKMQLAFRLQQIAAALE
839	2189	A	6872	1	1485	NKVTDL
1 637	2107	Α.	06/2	1	1465	RARRLALQCHVCVCALTPGEQSGRRLPGQT
		i			1	WLMFSCFCFSLQDNSFSSTTVTECDEDPVSLH
J	j				ĺ	EDQTDCSSLRDENNKENYPDAGALVEEHAPP
1		ĺ				SWEPQQQNVEATVLVDSVLRPSMGNFKSRKP
1						KSIFKAESGRSHGESQETEHVVSSQSECQVRA
1	1		[			GTPAHESPQNNAFKCQET\VRL\QPRIDQRTAT
}			i .	,	ļ	SPKDAFETR\QDLNEEEAAQVHGVKDPAPAS
			<b>i</b>			TQSVLA\DGTDSADPSPVHKDGQNEADSAPE
			ľ			DLHSVGTSRLLL/YHITDGDNPTAVRHGCSL/F
1				-		SGQSQRFNLDPESAPSPPSTQQFMMPRSSSRC
1						SCGDGKEPQTTTQLTKHIQSLKRKIRKFEEKFE
1						QEKKYRPSHGDKTSNPEVLKWMNDLAKGRK
			i 1		*	OI KEI KI KI GEEOCGADACEDDUR I GEOMEND
]			1			QLKELKLKLSEEQGSAPKGPPRNLLCEQPTVP
I I						RENGKPEAAGPEPSSSGEETPDAALTCLKERR
]	]					EQLPPQEDSKVTKQDKNLIKPLYDRYRIIKQIL
840	2190	_	COM		0054	STPSLIPTIVSQDTCMLLLCTDV
040	2190	A	6873	2	2054	FFRFYFSFIRLFAMSLADLTKTNIDEHFFGVAL
	i					ENNRRSAACKRSPGTGDFSRNSNASNKSVDY
1			) j			SRSQCSCGSLSSQYDYSEDFLCDCSEKAINRN
			!	,		YLKQPVVKEKEKKKYNVSKISQSKGQKEISV
						EKKHTWNASLFNSQIHMIAQRRDAMAHRILS
						ARLHKIKGLKNELADMHHKLEAILTENOFLK
1 1						QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV
1						KNLRQLLRKSQEKERTLSRKLRETDSQLLKT
						KDILQALQKLSEDKNLAEREELTHKLSIITTK
			i			MDANDKKIQSLEKQLRLNCRAFSRQLAIETR
1 ' 1			i i	Î		KTLAAQTATKTLQVEVKHLQQKLKEKDREL
						EIKNIYSHRILKNLHDTEDYPKVSSTKSVQAD
]			1			RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG
1 I						
	ĺ	1		İ	i	NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY
.				ļ		EDLSGEBKHLEVQILLENTGRQKDKKEDQEK
	,		<b> </b>	l		KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR
	1	- 1				EKFKRSMQRNGVDDT\LGKGTAPYTKGPLRQ
] ]						RRHYSFTEATENLHHGLPASGGPANAGNMR
į i	I	- 1		ł	ł	YSHSTGKHLSNREEMELEHS\DSGYEPSFGKS
ı l		l		ļ		SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD
].		ĺ				QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA
<u> </u>						FGDSKVTVVNSIKPSSPTEGKRKIII
841	2191	A	6874	3	2867	SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG
	i	1	ł	1	1	NAPAPGTPAASGWQPPTYHSGRAFSARYPRP
j [	1	I		l	Į.	SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA
]	1	J	l	ļ	į	DHAVRPLHGARGGQPPVPQQHVLERQVQLS
	ſ	ļ	ſ	į	F	QGQNVVIKVKPPSKSGSASASGAORGSLEEFE
		ļ			İ	DTPWSDQRPREGEGEPPRGQLQPSRPTRARG
[						TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP
, ,	į	j	J	l		DEDDD THOSE OF COMES OF FUND ACTION
} I	- 1			j	l	REPRRTVSESVIAVKASFPSSALPPRTGVALG
j l			1		l	RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP
	]	i	1	ı	i	SGSVGGPARPASGPRQAREASLVVTCRTNKF
, ,	- 1	}	. 1	}	}	RKNNYKWVAASSKSPRVARRALSPRVAAEN
		1	ŀ	ì	ĺ	VCKASAGMANKVEKPQLIADPEPKPRKPATS
1 1		ļ		l	1	SKPGSAPSKYKWKASSPSASSSSSFRWOSEAG
		ł		l	l	SKDHASQLSPVLSRSPSGD\RPALAHSGLKPLS
		İ		l	I	GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG
	İ	l		l	l	TSPAATAKSHLSLRRRQALRGKSSPVLKKTPN
						N11AAJV 166AUAJYAAAGALIOLIOLIOLIOLIOLIOLIOLIOLIOLIOLIOLIOLIOL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methioriine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  KGLVQVTKHRLCRLPPSRAHLPTKEASSLHA  VRTAPTSKVIKTRYRIVKKTPASPLSAPPFPLS  LPSWRARRLSLSRSLVLNRLRPVASGGGKAQ
842	2100		<b>4000</b>	606	2071	PGSPWWRSKGYRCIGGVLYKVSANKLSKTSG QPSDAGSRPLLRTGRLDPAGSCSRSLASRAVQ RSLAIIRQARQRREKREYCMYYNRFGRCNR GERCPYIHDPEKVAVCTRFVRGTCKKTDGTC PFSHHVSKEKMPVCSYFLKGICSNSNCPYSHV YVSRKAEVCSDFLKGYCPLGAKCKKKHTLLC PDFARRGACPRGAQCQLLHRTQKRHSRRAAT SPAPGPSDATARSRVSASHGPRKPSASQRPTR QTPSSAALTAAAVAAPPHCPGGSASPSSSKAS SSSSSSSSPPASLDHE\APSLQEAALAAACSNR LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKDS GKPLHIKPRL
	2192		6898	506	2071	WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH RNKFKVINVDDDGNELGSGIMELTDTELILYT RKRDSVKWHYLCLRRYGYDSNLFSFESGRRC QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFAA QNLPNGYPRYPSFGDASSHPSSRHPSVGSARL PSVGEESTHPLLVAEEQVHTYVNTTGVQEER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDAP SVNKLVYENINGLSIPSASGVRRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRRDCTPTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQTPTRRTELYAVIDIERTAAMSNLQKAL PRDDGTSRKTRHNSTVDLPL
843	2193	A	6919	2	663	AGRPGTTHASOKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS L/VFLGPGLYNN/GSSMCGAE\EPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF
844	2194	A	6928	902	366	HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM
845	2195	A	6939	1660	317	LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLIPHFKPWLHPEQSP LPSLALSYELSVQHADSYLENIDESAVAESREE RYMGGAGGEGYSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP RRVLQPGWTEPQ

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghttamic Acid.
nucl-	peptide	1100	in No.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
846	2196	A	6944	sequence 42	0670	nucleotide insertion
040	2190	^	0944	42	2672	RRKMAGCRGSLCCCCRWCCCCGERETRTPE ELTILGETQEEEDEILPRKDYESLDYDRCINDP
						YLEVLETMDNKKGRRYEAVKWMVVFAIGV
)						CTGLVGLFVDFFVRLFTQLKFGVVQTSVEECS
1 1						QKGCLALSLLELLGFNLTFVFLESLLGLIEPVE
						AGSGITEGKCYLYARQVPGLVRLPTLLWKAL
1						GVLLTVAAMLLI\GLGSPMIHSGSVVGAGLPQ
i i		i				FQSISLRKIQFNFPYFRSDRYGK\DKRDFVSAG
						AAAGVAAAFGAPIGGTLFSLEEGSSFWNQGL
1						TWKVLFCSMSATFTLNFFRSGIQFGSWGSFQL PGLLNFGEFKCSDSDKKCHLWTAMDLGFFV
						VMGVIGGLLGATFNCLNKRLAKYRMRNVHP
1						KPKLVRVLESLLVSLVTTVVVFVASMVLGEC
i						RQMSSSSQIGNDSFQLQVTEDVNSSIKTFFCP
1 1				İ		NDTYNDMATLFFNPQESAILQLFHQDGTFSPV
1						TLALFFVLYFLLACWTYGISVPSGLFVPSLLC
1 1	- 1			İ		GAAFGRLVANVLKSYIGLGHIYSGTFALIGAA
						AFLGGVVRMTISLTVILIEST\NEITYGLPIMVT LMVGKWTGDFFNKGI\YDIHVGLRGVPLLEW
1		·		i		ETEVEMDKLRASDIMEPNLTYVYPHTRIQSLV
						SILRTTVHHAFPVVTENRGNEKEFMKGNQLIS
1 1		· ·				NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL
1 1						RNMCDEHIASEEPAEKEDLLQQMLERRYTPY
	[			ĺ		PNLYPDQSPSEDWTMEERFRPLTFHGLILRSQ
l j		J			ļ	LVTLLVRGVCYSESQSSASQPRLSYAEMAED YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSPF
1.				·		TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE
	J			}		IVGITTRHNLTYEFLQARLRQHYQTI
847	2197	Α	6951	3	1994	NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER
1 1	j	1		ļ		LSSIEKIKQLREQVNDLFSRKFGEAIGVDFPVK
<b> </b> .		i				VPYRKITFNPGCVVIDGMPPGVVFKAPGYLEI
				1		SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST
			1			VGKRKIDQEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY
	ļ					MERMRDEKLHELKKGLRKYLLGLSDTECPE
	İ	ľ				QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR
1 1	İ	ŀ	1		. 1	EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE
	İ				Ì	NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK
1 1	i	ì	1	.	1	NFCINWSKLVSVASTGTPPMVDANNGLVTKL
		j	İ	l		KSRVATFCKGAELKSICCIIHPESLCAQ\KLKM
]	ł	ł	ł	ļ		DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI
1	l			l		DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM
		I		l	I	HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL
}		Į		1		WETHLTRNNLAHFPTLKLVSRNESDGLNYIP
1 1		ŀ				KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI
1 /	1	ł		ł		DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE
]		1			1	FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ
848	2198	A	6985	3	289	LFSIMKLSKTKYCSQLKDSQWDSVLHIAT SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK
			3,03	·	207	ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C
[ ]	1	ł	ł	ì	ł	CFNAINTKIPIQRLESYTRITNIQCPKEAVM
849	2199	A	6999	963	5	LDFLCHRDMGDNITSITEFLLLGFPVGPRIOM
1 1	ł	ł	ł		ł	LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP
i		l	}	ŀ		MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP
	ĺ	ď	ł	ŀ	1	AKPISFAGRMMQTFLFSTFAVTECLLLVVMS
}	j	Į			. [	YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT
	ł	ł		ł	ì	TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA
[			1	Į.	-	VLKLACADTHINENMVLAGAISGLVGPLSTIV VSYMCILCAILQIQSREVQRKAFCTCFSHLCVI
<b></b>		L				TO THE CHICAGO OF THE CHICAGO OF THE CANADA OF THE CHICAGO OF THE

NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  No: of peptide sequence  No: of peptide sequence  NO: of peptide sequence  No: of peptide sequen	
eotide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to la	
sequence    Sequence   Uence   09/496   Corresponding to last amino acid residue of peptide sequence   09/496   Corresponding to last amino acid residue of peptide sequence   M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion   GLFYGTAIIMYVGPRYGNPKEQKKYLLLF LFNPMLNPLICSLRNSEVKNTLKRVLGVE    1011   MGNDSVSYEYGDYSDLSDRPVDCLDGAC   DPLRVAPLPLYAAIFLVGVPGNAMVAWV	
uence  914  ng to first amino acid residue of peptide residue of peptide sequence  914  ng to first amino acid residue of peptide sequence  914  ng to first amino acid residue of peptide sequence  925  934  125  126  127  128  138  138  138  138  138  138  138	
amino acid residue of peptide sequence	
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  GLFYGTAIIMYVGPRYGNPKEQKKYLLLF LFNPMLNPLICSLRNSEVKNTLKRVLGVEI  850 2200 A 7001 1 1011 MGNDSVSYEYGDYSDLSDRPVDCLDGAC DPLRVAPLPLYAAIFLVGVPGNAMVAWV	
peptide sequence /possible nucleotide deletion, possible nucleotide insertion /possible nucleotide insertion / GLFYGTAIIMYVGPRYGNPKEQKKYLLLF LFNPMLNPLICSLRNSEVKNTLKRVLGVEI / LFNPMLNPLICSLRNSEVKNTLKRVLGVEI / DPLRVAPLPLYAAIFLVGVPGNAMVAWV	
Sequence nucleotide insertion  GLFYGTAIIMYVGPRYGNPKEQKKYLLLF LFNPMLNPLICSLRNSEVKNTLKRVLGVEI  850 2200 A 7001 1 1011 MGNDSVSYEYGDYSDLSDRPVDCLDGAC DPLRVAPLPLYAAIFLVGVPGNAMVAWV	
850 2200 A 7001 1 1011 MGNDSVSYEYGDYSDLSDRPVDCLDGAC DPLRVAPLPLYAAIFLVGVPGNAMVAWV	
850 2200 A 7001 1 I011 MGNDSVSYEYGDYSDLSDRPVDCLDGAC DPLRVAPLPLYAAIFLVGVPGNAMVAWV	LIC
850 2200 A 7001 1 1011 MGNDSVSYEYGDYSDLSDRPVDCLDGAC DPLRVAPLPLYAAIFLVGVPGNAMVAWV	
DPLRVAPLPLYAAIFLVGVPGNAMVAWV	
PIARGGHWPYGAVGCRALPSIILLTMYAS	
LAALSADLCFLALGPAW\CLRFS\GACGVC	
CGAAWTLALLLTVPSAIYRRLHQEHFPAR	
CVVDYGGSSSTENAVTAIRFLFGFLGPLVA	VA
SCHSALLCWAARRCRPLGTAIVVGFFVCV	
YHLLGLVLTVAAPNSALLARALRAEPLIV	
ALAHSCLNPMLFLYFGRAQLRRSLPAACE	
ALRESQGQDESVDSKKSTSHDLVSEMEV	
851 2201 A 7011 1 2310 AAASPLRMSRKGPRAEVCADCSAPDPGW	ASI
SRGVLVCDECCSVHRSLGRHISIVKHLRHS	A
WPPTLLQMVHTLASNGANSIWEHSLLDPA	QV
QSGPALKQTPKDKV\HPIKSEFIRAKYQML	
VHKLPCRDDDGVTAKDLSKQLHSSVRTG	NLE
TCLRLLSLGAQANFFHPEKGTTPLHVAAK	AG
QTLQAELLVVYGADPGSPDVNGRTPIDYA	
AGHHELAERLVECQYELTDRLAFYLCGRE	
HKNGHYIIPQMADSLDLSELAKAAKKKLQ	
SNRLFEELAMDVYDEVDRRENDAVWLAT	
HSTLVTERSAVPFLPVNPEYSATRNQGRQI	
ARFNAREFATLIIDILSEAKRRQQGKSLSSP	TD
NLELSLRSQSDLDDQHDYDSVASDEDTDQ	
LRSTGATRSNRARSMDSSDLSDGAVTLQE FLYVALATSFAVVOOLAGVISHER SDELED	
ELKKALATSEAKVQQLMKVNSSLSDELRF REIHKLQAENLQLRQPPGPVPTPPLPSERAI	
TPMAPGGSTHRRDRQAFSMYEPGSALKPF	
PPGDELTTRLQPFHSTELEDDAIYSVHVPA	
YRIRKGVSASAVPFTPSSPLLSCSQEGSRHI	
LSRHGSGADSDYENTOSGDPLLGLEGKRF	
LGKEEDFHPELESLDGDLDPGLPSTEDVILI	
EQVTKNIQELLRAAQEFKHDSFVPCSEKIH	I.A
VTEMASLFPKRPALEPVRSSLRLLNASAYR	
SECRKTVPPEPGAPVDFQLLTQQVIQCAYE	
KAAKQLVITTTREKKQ	- 1
852 2202 A 7016 484 1777 RISKIQVYYSTGYSSRKMNPTLGLAIFLAVI	T
TVKGLLKPSFSPRNYKALSEVQGWKQRM	AA
KELARQNMDLGFKLLKKLAFYNPGRNIFL	SP
LSISTAFSMLCLGAQDSTLDEIKQGFNFRKI	
EKDLHEGFHYIIHELTQKTQDLKLSIGNTLI	
QRLQPQRKFLEDAKNFYSAETILTNFQNLE	
AQKQINDFI/ESKTHGKINNLIENIDPGTVMI	
ANYIFFRARWKHEFDPNVTKEEDFFLEKNS	
VKVPMMFRSGIYQVGYDDKLSCTILEIPYQ	K
NITAIFILPDEGKLKHLEKGLQVDTFSRWK	TL
LSRRVVDVSVPRLHMTGTFDLKKTLSYIGV	
KIFEEHGDLTKIAPHRSLKVGEAVNKAELK	
DERGTEGAAGTGAQTLPMETPLVVKIDKP	YL
LLIYSEKIPSVLFLGKIVNPIGK   853   2203   A   7017   1   3293   MTHACNPSTI GGOGPPITPSHGPPPP SEPCE	
	V
ARHVAAGAGHENKHGGSRRFPAGVAPRR	AM
ANVSKKVSWSGRDRDDEEAAPLLRRTARP	G
GGTPLLNGAGPGAARQSPRSALFRVGHMS	
ELDDELLEP\DMDPPHPFPKEIPHNEKLLSLI	
ESLDYDNSENQLFLEEERRINHTAFRTVEIK	
WVICALIGILTGLVACFIDIVVENLAGLKYR	VI
KGSILPNIDKFTEKGGLSFSLLLWATLNAAI	·V

SEQ ID NO: of nucl- cotide	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid residue of	corresponding to last amino acid residue of peptide sequence	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
						LVGSVIVAFIEPVAAGSGIPQIKCFLNGVKIPH VVRLKTLVIKVSGVILSVVGGLAVGKEGPMI HSGSVIAAGISQGRSTSLKRDFKIFEYFRRDTE
						KRDFVSAGAAAGVSAAFGAPVGGVLFSLEEG ASFWNQFLTWRIFFASMISTFTLNFVLSIYHG NMWDLSSPGLINFGRFDSEKMAYTIHEIPVFI
						AMGVVGGVLGAVFNALNYWLTMFRIRYIHR PCLQVIEAVLVAAVTATVAFVLIYSSRDCQPL QGGSM\$YPLQLFCADGEYNSMAAAFFNTPBK SVVSLFHDPPGSYNPLTLGLFTLVYFFLACWT
						YGLTVSAGVFIPSLLIGAAWGRLFGISLSYLTG AAIWADPGKYALMGAAAQLGGIVRMTLSLT VIMMEATSNVTYGFPIMLVLMTAKIVGDVFIE
						GLYDMHIQLQSVPFLHWEAPVTSHSLTAREV MSTPVTCLRRREKVGVIVDVLSDTASNHNGF PVVEHADDTQPARLQGLILRSQLIVLLKHKVF
				·		VERSNLGLVQRRLRLKDFRDAYPRFPPIQSIH VSQDERECTMDLSEFMNPSPYTVPQEASLPR VFKLFRALGLRHLVVVDNRNQVVGLVTRKD
						LARYRLGKRGLEELSLAQTGPKAQATAEGRV AGAAQQPCQLRAVTLEDLGLLLAGGLASPEP LSLEELSERYESSHPTSTASVPEQDTAKHWNQ LEQWVVELQAEVACLREHKQRCERATRSLL
						RELLQVRARVQLQGSELRQLQQEARPAAQAP EKEAPEFSGLQNQMQALDKRLVEVREALTRL RRRQVQQEAERRGAEQEAGLRLAKLTDLLQ
854	2204	A	7037	139	2604	QEEQGREVACGALQKNQEDSSRRVDLEVAR M
034		A	7037	139	2004	AGTWEPRPYDQAKETGAPGSQPPVPPMELRP WLLWVVAATGTLVLLAADAQGQKVFTNTW AVRIPGGPAVANSVARKHGFLNLGQIFGDYY
						HFWHRGVTKRSLSPHRPRHSRLQREPQVQWL EQQVAKRRTKRDVYQEPIDPKFPQQWYL\SG VTQ\RDLMVKAAWAQGYTGHGIVVSILDDGI EKNHPDLAGNYDPGASFDVNDQDPDPQPRY
	•					TOMNDNRHGTRCAGEVAAVANNGVCGVGV AYNARIGGVRMLDGEVTDAVEARSLGLNPN HIHIYSASWGPEDDGKTVDGPARLAEEAFFR
					·	GVSQGRGGLGSIFVWASGNGGREHDSCNCD GYTNSIYTLSISSATQFGNVPWYSEACSSTLA TTYSSGNQNEKQIVTTDLRQKCTESHTGTSAS
	ı					APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTBPKDI
						GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN
						TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSROS
						QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD
855	2205	A	7058	3	1441	SEEDEGRGERTAFIKDQSAL  QRPASQLLAPFAAEALPGAPRAAMAQHFSLA
				:		ACDVVGFDLDHTLCRYNLPESAPLIYNSFAQF LVKEKGYDKELLNVTPEDWDFCCKGLALDL EDGNFLKLANNGTVLRASHGTKMMTPEVLA
L		L	l	L	L	EAYGKKEWKHFLSDTGMACRSGKYYFYDN

SEQ ID NO: of nucl- ectide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide Insertion
						KDIVAAIQHNYKMSAFKENCGIYFPEIKRDPG RYLHSRPESVKKWLRQLKNAGKILLLITSSHS DYCRLLCAIYILGNDFTDLFDIVITNALKPGFP SHLPSQRPFRTLENDEEQEALPSLDKPGWYSQ GNAVHLYELLKKMTGKPEPKVVYFGDSMHS DIFPARHYSNWETVLILEELRGDEGTRSQRPE ESEPLEKKGKYEGPKAKPLNTSSKKWGSFFII DSVLGLENTEDSLVYTWSCKRISTYSTIAIPSI EAIAELPLDYKFTRFSSSNSKTAGYYPNPPLV LSSDETLISK
856	2206	A	7082	396	1635	SSPSVFEFEHAVQPVFTMEFLKTCVLRRNACT AVCFWRSKVVQKPSVRRISTTSPRSTVMPAW VIDKYGKNEVLRFTQNMMMPIIHYPNEVIVK VHAASVNPIDVNMRSGYGATALNMKRDPLH VKIKGEEFPLTLGRDVSGVVMECGLDVKYFK PGDEVWAAVPPWKQGTLSEFVVVSGNEVSH KPKSLTHTQAASLPYVALTAWSAINKVGGLN DKNCTGKRVLILGASGGVGTFAIQVMKAWD AHVTAVCSQDASELVRKLGADDVIDYKSGSV EEQLKSLKPFDFILDNVGGSTETWAPDFLKK WSGATYVTLVTPFLLNMDRLGIADGMLQTG VTVGSKALKHFWKGVHYRWAFFMASGPCL DDIAELVDAGKIRPVIEQTFPFSKVPEAFLKV ERGHARGKTVINVV
857	2207	A	7088	320	2417	LRRRKMTPQSLLQTTLFLLSLLFLVQGAHGR GHREDFRFCSQRNQTHRSSLHYKPTPDLRISIE NSEEALTVHAPFPAAHPASRSFPDPRGLYHFC LYWNRHAGRLHLLYGKRDFLLSDKASSLLCF QHQEESLAQGPPLLATSVTSWWSPQNISLPSA ASFTFSFHSPPHTGAHNASVDMCELKRDLQL LSQFLKHPQKASRRPSAAPASQQLQSLESKLT SVRFMGDMGSFEEDRINATVWKLQPTAGLQ DLHIHSRQEEEQSEIMEYSVLLPRTLFQRTKG RSGEAEKRLLLVDFSSQALFQDKNSSQVLGE KVLGIVVQNTKVANLTEPVVLTFQHQLQPKN VTLQCVFWVEDPTLSSPGHWSSAGCETVRRE TQTSCFCNHLTYFAVLMVSSVEVDAVHKHY LSLLSYVGCVVSALACLVTIAAYLCSRVPLPC RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPV ALTGSEAGCRASAIFLHFSLLTCLSWMGLEG YNLYRLVVEVFGTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPILAVHRTPEGVTYPS MCWIRDSLVSYTTNLGLFSLVFLFNMAMLAT MVVQILRLRPHTQKWSHVLTLLCLSLVLGLP WALIFFSFASGTFQLVVLYLFSITSFQGFLIFI WYWSMRLQARGGPSPLKSNSDSARLPISSGS TSSSRI
858	2208	A	7091	185	415	DAGAVKSSDTNIWFRGMCDDKKGHRCPS*G QPQHFHVAFHTEAEGAMFYFRLHVIHRVMQS QQQLFPSTLFSWLLE
859	2209	A	7136	3	302	FFFWRQSLALLPRLECSGATGAHCNLHFPGSS DCPTSAS*IAGITGACYHAWLLFVFLAETGFH HVGQGGLELLTSSDPSGSASQSAGITGVSHCT WPI
860	2210	A	7156	23	591	ALSTETRTPDMRRLLLVTSLVVVLLWEAGAV PAPKVPIKMQVKHWPSEQDPEKAWGARVVB PPEKDDQLVVLFPVQKPKLLTTBEKPRGQGR GPILPGTKAWMETEDTLGRVLSPEPDHDSLY

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, B=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isolcucine, K=Lysine, L=Lcucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide inscrtion
						HPPPEEDQGEERPRLWVMPNHQVLLGPEEDQ DHIYHPQ*GSRGHHCPRPVPRPRLLGLGPSLP CPS
861	2211	A	7161	1220	1003	NYVCTIAF*EKKMGF*LSLSCLVLLFVLFLDCI LTTTTRIMFHCTYLFASVCLSLLNTLLSPNCL KSAMILQ
862	2212	A	7211	665	847	LKYYHITMGIYKTGKKVIL*KSSMSNRFSVIF YKNIQKLSFSNYVYHQNYVFSSDWSYDF
863	2213	A	7212	924	1273	HGSSCALGDLAPG*LPSGPVLSSPAVRL*RKP LVWDSPSCLPATGPT*GLVLVLGGPDCT*WA RGQHEHKRMRAP*SCRVTVNLAKKKKKTDQ CIKPNYQSPPKECDYNILANSVA
864	2214	A	7214	845	1619	SDKGGKKADRKNHI.RHAFPLI.PHRVRERLH DPKVPVDADHVQGQDPGRAAHDIHGEDVTE KVSKDPLAPDEVGDTDEGHDRHGHREVGQR HGHDQEEVAYEERACEGGKFATVEVTDKPV DEALREAMPKVAKYAGGTNDKGIGMGMTV PISFAVFPNEDGSLQKKLKVWFRIPNQFQSDP PAPSDKSVKIEEREGITVYSMQFGGYAKEAD YVAQATRI.RAALEGTATYRGDIYFCTGYDPP MKPYGRRNEIWLLKT
865	2215	A	7246	559	682	RRLGAVAHAYTSSTLGGRGGWIT*GQELQTS LANMAKPRLY
866	2216	A	7257	641	1310	TCTYKYLMGWIRGRRSRHSWEMSEFHNYNL DLKKSDFSTRWQKQRCPVVKSKCRENASPFF FCCFIAVAMGIRFIIMVAIWSAVFLNSLFNQEV QIPLTESYCOPCPKNWICYKNNCYQFFDESKN WYESQASCMSQNASLLKVYSKEDQDLLKLV KSYHWMGLVHIPTNGSWQWEDGSILSPNLLT IIEMQKGDCALYASSFKGYIENCSTPNTYICM QRIV
867	2217	A	7288	151	396	SIKIIEAFGSNGPDFWFFRYWSP*LFRQQVVFI MPFFQTLWLMNANRFCSIFTTTNVANNCWW TPYHCWLSVVVCRCESHGI
868	2218	A	7298	3	272	PDTVIGGRGSGGKEFGRWVLW*VFE*RLGTP KGSCPAGGSRMVSESD*EGRGC*ASYPCAC* AGS*WR*GSRPAGRGTPPRSLSHARPP
869	2219	A	7332	1223	332	PRRDAEDRDESCLNPAFPIGLLHPNSVNSMAR FLTLCTWLLLLGPGLLATVRAECSQDCATCS YRLVRPADINFLACVMECEGKLPSLKIWETC KELLQLSKPELPQDGTSTLRENSKPEESHILLA KRYGGFMKRYGGFMKKMDELYPMEPEEEA NGSEILAKRYGGFMKKDAEEDDSLANSDLL KELLETGDNRERSHHQDGSDNEEEVSKRYGG FMRGLKRSPQLKEKAKELQKRYGGFMRRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF
870	2220	A	7382	216	1018	EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENIT RAEALTEPLNA
871	2221	A	7403	3	393	SCAMCSGLL*LLLPIWLSWILGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR ALRGAALPGESEAGDPESLRSSVNADWIQYS

NO: of   No: of   N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
muclectide   seq							
Sequence				1			
Sequence	1						
1914   18	,		i	1			
aniilos acid recidue of peptide requence   peptide sequence   peptid							
residue of peptide sequence		Ì					
Popesible nucleotide deletion, \( =\) possible nucleotide disention   DLWEAEVSTRCEAGFCQECFRTPGNQEKDG		j	1	ļ			
DI.WEARVSTPRCEAGFCGECFRTPGNQEKDO   PPIC	<b>!</b> '	ì					
872   2222   A   7413   1061   359	·	<del></del>					
\$72   \$222   A   \$7413   \$1061   \$359	1 .	J	ļ	)	ļ		
### PGGS*PQATI.HI.DRAMEVASSTREIIQVIKYEK CGLIKPCPANYTAPKICSGAANVVGPTMCFT RMMSPYKNIN/GRCIAIIALVMGTTGAVI.QQ KAPDMYSGIVMEIL-VERLERGALVI.VAS YDDPGTKMMDSSRILFISIJGSSYAKQI.GFRD SWVFIGAKDLRGKSPFEQFLKEQPQTQNKYE GWPEILEMEGGMPKPP  ### 2224 A 7429 2242 2394 ILKCAGHIGGSCI*SQHFGRLRWEDRLRI.GVQ DHPGQHCEITSLIKBRKLP ### 2224 A 7468 146 894 PCTSCVI.WATLHI.PASTIKAPQAECAMISTIC GWPUTLLRWELLGMILETYGFDSVILAFGN LLFITGLSLIGLRKTFWFPQRHKILKGTSLIF GWPUTLLRWELLGMILETYGFPLSVILAFGN LLFITGLSLIGLRKTFWFPQRHKILKGTSLIF AFGFLGNIVCNIPFI.GALLFRILQGTSSMY-KEI MSSIALDHWILKGAKREEWEPPOSSPALTHSP TYPGPPQVQERNGABQLTSNPQVDSRGQQ AEMQTPRRI.GWWYHITLIT.LWEEK #### 2225 A 7498 91 251 GEEPVPTWLQDEAGQWLIGPVAQPWGWPG SERHEP*HGGVLRGJRAPPGSFALTHSP TYPGPPQVQERNGABQLTSNPQVDSRGQQ AEMQTPRRI.GWGWYHITLIT.LWEEK ### 357 2225 A 7566 2 940 GCAPPTREFYDEFGGRGAAPWVALVARGGC TIFKDKVLVAARRASAVULYNEERYGRITUM ### 358AGTIGNIVVMINSTYRGRELLEVQKGIPV TIMITIOVGTRHVQEPISGQSVVFVAIAFITMMI SLAWLIFYYTQRFLYTOSGIGGSSMREETKKWI GQLLHTYKHGBKGIDVDAENCAVCENFYK KURIRLPCHIHERICOPPULDHRICO	872	2222	A	7413	1061	350	
			1	' ' ' '		1	
RAMMSPYKNNVGRGLNIALVNGTTGALQG KAPDMYSGDVMILVKFKERIPGGALVLVAS YDDPGTKMDESRKLFSDLGSSYAKQLGFGA SVPUGAZDLRGKSPPEGREVEGORYORKFE GWPELLEMGCMPPKPF   ROME		ļ	ĺ				
RAFDMYSGDVMHLVKFLKEITGGALVLVAS   YDDPGTKMNDESKLFSDLGSSYAKQLGFED    SWYFIGAKDLRGKSPFEQFLKEQPQTQNKYE    GWPLLEMEGGMPPKPF    SWYFIGAKDLRGKSPFEQFLKEQPQTQNKYE    GWPLLEMEGGMPFKPF    SP4   2224   A 7468	ļ ·	İ	l		}	İ	
STD	1						
SWYFIGAKDLRGKSFFEQFLKEQPQTONKYE   GWPPLLEMEGCMPPKFF				1			:
873   2223   A   7429   2242   2394   ILKZAGHGGGSC 1*SQHFGIRRUDDIRIT.GVQ DHFQQHCETPSILKIRRIL.P     874   2224   A   7468   146   894   PCTSCVIWATILEPASTRKAPQAECGMISTTE     WQKIGVGITGGIFFILFGTLYFDSVILLAFGN   LLFLTGLSLIGLRKTEWFFFGRHLK.KGTSFLL     GGVVIVILLRWHJLGMFLETTGFSLFKGFFYC     A   7498   91   251   GEKPVTWLQDEAGQWLLGFVAQPWGWFG     SERHP*HGGVIFFLGKGFFYC     875   2225   A   7498   91   251   GEKPVTWLQDEAGQWLLGFVAQPWGWFG     876   2226   A   7544   403   587   YSCLCFLFKHTTSFRSVSIWLGTVVHAYNPN     1.GQQGGMA*GQFTSLGNTVRPLVARGGG     877   2227   A   7566   2   940   GCAPDTRFFYBEPGGRGAFWALVARGGG     1.GQQGMA*GQFTSLGNTVRPLVARGGG     1.GQQGMA*GQFTASLGNTVRPLVARGGG     1.GQQGMA*GQFTASLGNTVRPLVARGGG     1.GQQGMA*GQFTASLGNTVRPLVARGGG     1.GQQGMA*GQFTASLGNTVRPLVARGGG     1.GQQGMA*GQFTASLGNTVRPLVARGGG     1.GQQGGMA*GQFTASLGNTVRPLVARGGG     1.GQQGGMA*GQGGSTASLGGSASLGGASLTASLGGGG     1.GQQGGGGTASLGGGASLGGASLGGASLGGASLGGASLGGASLG	1		ļ			!	
873   2223   A   7429   2242   2394		1		[			
B74   2224   A   7468   146   894   PCTSCVLWATLHEPASTRKAPQAECGMISTIE   WQKIGVGTTGFGIFFILYTEDSVLLAFGN   LLFLTGLSLIGLRKTFWFFPQRHKLGTSFIL   GGVVVILLAFGN   LLFLTGLSLIGLRKTFWFFPQRHKLGTSFIL   GGVVVILLAFGN   LLFLTGLSLIGLRKTFWFFPQRHKLGTSFIL   GGVVVILLAFGN   LLFLTGLSLIGLRKTFWFFPQRHKLGTSFIL   MSSINLDHWLGGAKEEWEPPPQSPALTHSP   TYPOPPQVQKERNGAEQLTSNPQVDSRGQQE   AEMQTPRRLGWGWYHTLTVJ.WEEK     875   2225   A   7498   91   251   GEKPVFTWLQDEAGGWLLGFVAQPWGWPG   SERHEP*HGGVLFRLGPSAPPGKL   SERHEP*HGGVLFRLGPSAPPGKL   SERHEP*HGGVLFRLGPSAPPGKL   SERHEP*HGGVLFRLGPSAPPGKL   GGAPTTRFPVPEFGGRAPWVALVARGG   TFLDKVLVAARNASAVVLVMEERYGNTILP   MSHAGTGNIVVIMISYPKGRELVK   UVKALPV   TMTIGVGTRHVQEFISQGSQSVVFVAIAFITMMI   SLAWLIFYYTGRFLYTGSQGSQSVTFVAIAFITMMI   SLAWLIFYYTGRFLYTGSQGSQSVTFVAIAFITMMI   SLAWLIFYYTGRFLYTGSQGSQSVTFVAIAFITMMI   SLAWLIFYYTGRFLYTGSQGSQSVTFVAIAFITMMI   SLAWLIFYYTGRFLYTGSQGSQSVTFVAIAFITMMI   SLAWLIFYYTGRFLYTGSQGSQSVTFVAIAFITMMI   SLAWLIFYYTGRFLYTGSQGSQSVTFVAIAFITMMI   SLAWLIFYYTGRFLYTGSQGSQSVTFVAIAFITMMI   SLAWLIFYYTGRFLYTGSQGSQSVTFVAIAFITMMI   SLAWLIFYYTGRFLYTGSGGSQSVTFVAIAFITMMI   SLAWLIFYYTGRFLYTGSGGSQSVTFVAIAFITMMI   SLAWLIFYYTGRFLYTGSGGSQSTRETKKY   UVKALGVGFPGDDVQEMPAPESPPGRDPAA   NLSLALPDDDGSDESSPSASPAGESPQCDPSF   KGDAGENTALLAGGRSDSRHGCKSG   GKSELLKSGSKSTLKHWTESKDLSISRILS   QTTRGKENDTDLDLRYDTYFEYGDLWDW   LRNSTDLQFPRFAKRRFPKTKTGFKKKMFGW   GDFHSNIKTVKINLLTGKIVDHGNGTFSVYF   RHNSTGGGNVSVSLVYPTKIVEPDLAQQTYD   AKDSKSFNCRIEYEKVDKATKNTILCNYDPSK   TCYGEQTQSHVSVSLCSKFFKVCTYISFYSTD   YKLVQKVCPDYNYHSDTYPFSG   GGSGRSDRGSGQODSLYPVGVLDKQVPDTS   VQETDRILVEKRCWDIALGFLKQPDTS   VQETDRILVEKRCWDIALGFLKQNPDTS   VQETDRILVEKRCWDIALGFLKQNPDTS   VQETDRILVEKRCWDIALGFLKQNPDTS   VQETDRILVEKRCWDIALGFLKQNPDTS   VQETDRILVEKRCWDIALGFLKQNPDTS   VQETDRILVEKRCWDIALGFLKANTGNTL   LNVALABLLIFCLFPFRINTHNTTTSVSSWPYSSERMRFIT   NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT   SYSVIFIVGLVGNIFLANTYLINGNIKYLLGVLLICHTMI   VQRAFITHMI   VGGRATTTWHNONKWTLGVLLICKVVGTLFYMMWISIILLGFISLGPG   QQRAITTKOSSIYVCCVVGTLFYMMWISIILLGFISLGRG   QQRAITTKOSSIYVCCVVGTLFYMMINGNIV   VCKVVGTLFYMMMXISILL	873	2223	A	7429	2242	2394	
\$74	1						
### WQKIGVGITGFGIFFLYGTIL1YFDSYLLAFGN LLFLTGLSLIGIGKTIFFFFGHKKGTSFLL GGVVIVLLRWPLLGMTLEHTYGFFSLFKGFFV AFGILGHVCNIFFLGALFRLQGTSSMY*KTE MSSLNLDHWLGGAKEEWEPPPQSPALTISP TYPGPPQVQKERNGAEQLTSNPQVDSRGCQE AEMQTPRALGWGWYHTLT-LYLWEEK ### WASHLANDHUGAKEEWEPPPQSPALTISP TYPGPPQVQKERNGAEQLTSNPQVDSRGCQE #### WASHLAFTGHTSFKNSVHWLGTVVHAYNPN ILGQQGWIA*QQEFKTSLGNTVFRCLYK #### USZZZZ A 7566 2 940 GCAPTIFFYPEFGGRAPWVALVARGGC GCAPTIFFYPEFGGRAPWVALVARGGC TFKDKVLVAARNASAVVLYNEREYGNTILP MSHAGTGHVVUMISYRFAELLELVQKGIPV TMTIGVGTRHVQEFISQQSVYPVAIAFITMMI SLAWLIFYTIGRFLYTGSGGSSIRKETKKVI QCILLHTVKHGEKGIDVDAENCAVCEBFKV KDIRLIPCKHIFHRICDPWLLDHRTCPMCKL DVIKALGYWGEFGDVQEWPAPESPFGRPPA NLSLALPDDDGSDESSPFSASPAGSBPQCDPSF KGDAGENTALLEAGRSDSRIGGFIS ####################################	874	2224	Α	7468	146	894	
		-	1				
GGVVIVILRWPILGMELETYGFFSLFKGFFPY							
### ### ##############################		ļ			·		
MSSLNLDHWLKGAKREEWEPPPGSPALTHSP			i	i i			
TYPGPPQVQKERNGABQLTSNPQVDSRGCQE	!		ł				MSSLNLDHWLKGAKREEWEPPPOSPALTHSP
AEMQTPRILGWGWYHTLTL/LWEEK		100					
875   2225   A   7498   91   251   GEKPVPTWLQDEAGQWLLGFVAQPWGWPG     876   2226   A   7544   403   387   YSCLCELFRHISFKNSVHWLGTVVHAYNPN     877   2227   A   7566   2   940   GCAPDTRFFVPEPGGRGAPWALVARGGC							
SERHEP*HGÖVLFRLĞPSAPPGKL     STA   2226   A   7544   403   587   YSCLCFLFKHITSFKNSVHIWLGTVVHAYNPN     LIGGQGGWIA-9GQERKTSLGNTVRPCLYK     STA   2227   A   7566   2   940   GCAPDTRFFVPEPGGRGAAPWVALVARGGC     TFKDKVLVAARRNASAVVLYNEERYGNITLP     MSHAGTGNIVVMISYPKGREILELVQKGIPV     TMTIGVGTRHVQEFISGQSVVFVAHAFITMMI     SLAWLFYY1QEFISGQSVVFVAHAFITMMI     SLAWLFYY1QEFISGQSVVFVAHAFITMMI     SLAWLFYY1QEFISGQSVVFVAHAFITMMI     SLAWLFYY1QEFISGQSVVFVAHAFITMMI     SLAWLFYY1QEFISGQSVVFVAHAFITMMI     SLAWLFYY1QEFISGQSVVFVAHAFITMKK     GQLLLHTVKHGEKGIDVDAENCAVCIENFKV     KDIRRLPCKHIFHRICIDPWLLDHRTCPMCKL     DVIKALGYWGEPGDVQEMPAFSPPGRPPA     NLSALPDDDGSDESSPSASPAESEPQCDPSF     KGDAGENTALLEAGRADRHGGFIS     RGRMQAACWYVLFLLQPTVYLVTCANLTNG     GKSELLKSGSKSTLKHWTESSKDLSISRLLS     QTFRGKENDTDLDLRYDTFEPYSEQDLWDW     LRNSTDLQEPRPRAKRPIVKTGKFKKMFGW     GFFSNIKTVKLNLLITGKIVDHGNGTFSVYF     RINSTGQGNVSVSLVPPTKVEFDLAQQTVID     AKDSKSFNCRIEVEKVDKATKNTLCTYDPSK     TCYQEQTQSHWUCSKPFKVICTVISFYSTD     YKLVQKVCPDYNYHSDTPYFPSG     S80   2229   A   7605   479   391   TESWKLKWWSPTCLDQLNGSAPGNVFHG     GGSRGRSDRGSQQDSLYPVGYLDKQVPDTS     VQETDRILVEKRCWDIALGPLKQIPMLFIMMY     MAGNTISIEPTIMMVCMMAWRIQALMAISAT     FKMLESSSQKFLQGLVYLIGNLMGLALAVYK     CQSMGLLPTHASDWLAFIEPFERMEFSGGGL     LL     LL     LL     SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT     NIBDQPPQNFSATPNVTTCMDEKLLSTVLTT     NIBDQPPQNFSATPNVTTCMDEKLLSTVLTT     SYSVEIVGLVGNIALYVFLGHRKRNSIQYT     LNVAIADLLLIFCLFFRIMYHINQNKWTLGVIL     CKVVGTTTFMNMYISILLGFISLDRYKINRSI     CKVVGTTTFMNMYISILLGFISLDRYKINRSI     CKVVGTTTFMNMYISILLGFISLDRYKINRSI     CKVVGTTTFMNMYISILLGFISLDRYKINRSI     CKVVGTTYFMNMYISILLGFISLDRYKINRSI     CKVVGTTYFWNMYISILLGFISLDRYKINRSI     CKVVGTTYFMNMYISILLGFISLDRYKINRSI     CKVVGTTYFMNMYISILLGFISLDRYKINRSI     CKVVGTTYFMNMYISILLGFISLDRYKINRSI     CKVVGTTYFWNMYISILLGFISLDRYKINRSI     CKVVGTTYFMNMYISILLGFISLDRYKINRSI     CKVVGTTYFMNMYISILLGFISLDRYKINGFI     CKVVGTTYFMMYISILLGFISLDRYKINGFI     CKVVGTTYFMMYTGMLALGGFLTMIL	875	2225	Α	7498	91	251	
876   2226   A   7544   403   587   YSCLCFLFKIITISFKNSVHIWLGTVVHAYNPN   ILGQQGGWL*GQGFKTSLGNTVRPCLYK   877   2227   A   7566   2   940   GCAPDTRFFVPEPGGRGAAPWALVARGGC   TFKDKVLVAARRNASAVVLYNEERYGNTTLP   MSHAGTONIVVIMISYPKGREILELVQKGIPV   TMTIGYGTRHVQEFISGQSVVFVAIAFITIMMI   SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI   GQLLLHTVKHGEKGIDVDAENCAVCIENFKV   KDIIRILPCKHIFHRICIDPWLLDHRTCPMCKL   DVIKALGYWGEPGDVQEMPAPESPPGRPAA   NLSLALPDDDGSDESPPSASPAESPQCDPSF   KGDAGENTALLEAGRSDSRHGGFIS   RGRMQAACWYVLFLLQPTVYLVTCANLTNG   GKSELLKSGSSKSTLKHIWTESSKDLSISRILS   QTFRGKENDTDLDLRYDTYPEYSEQDL WDW   LRNSTDLQEFPPRAKRRPIVKTGKFKKMPGW   GDPHSNIKTVKLINLLITIGKIVDHGNGTFSVYF   RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID   AKDSKSFNCRIEYEKVATKLCNYDPSK   TCYQEQTQSHVSWLCSKPFKVICTYJSFYSTD   YKLVQKVCPDYNYHSDTPYPFSG   GGSRGRSDRGSQQDSLYPGYJLDKQVPDTS   VQETORILVEKRCWDIALGFLKQIPMILFINY   MAGNTISIPPTMMVCMMAWRPIQALMAISAT   FKMLESSSQKFLQGLVYLIGHLMGLALAVYK   CQSMGLLPTHASDWLAFIEPPERMEFSGGGL   LL   LL   SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT   NHSDQPPQNFSATPNUCMMAWRPIQALMAISAT   FKMLESSSQKFLQGLVYLIGHLMGLALAVYK   CQSMGLLPTHASDWLAFIEPPERMEFSGGGL   LL   LL   SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT   NHSDQPPQNFSATPNUCMMAWRPIQALMAISAT   FKMLESSSQKFLQGLVYLIGHLMGLALAVYK   CQSMGLLPTHASDWLAFIEPPERMEFSGGGL   LL   LT   LNVAIADLLLFCLPFRIMYHINQNKWILGVIL   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINGST   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINRST   CKVVGTLFYMMYISILGFISLDRYKKINGST   CKVVGTLFYMMYISILGFISLDRYKINGST   CKVVGTLFYMMYISILGFISLDRYKGNSGOYL   CKVVGTLFYMMYISILGFISLDRYKINGST   CKVVGTLFYMMYISILGFISLDRYKINGST   CKVVGTLFYMMYISDRYMI   CKVVGTLFYMMYI CKVVGTLYWMALALGGFLTMILL							
ILGGQGWIA*GQEFKTSLGNTVRPCLYK   S77   2227   A   7566   2   940   GCAPDTRFFVPEPGGRGAPWVALVARGGC   TFKDKVLVAARRNASAVULYNEERYGNTTLP   MSHAGTGNTVVIMISYPKGREILELVQKQIPV   TMTIGVGTRHVQEFISGQSVVPVAIAFITIMMI   SLAWLIFYJQRFLYTGSQIGSQSHRKETKKVI   GQLLLHTVKHGBKGGDVDAENCAVCIENFKV   KDIDRIPCKHFHRICDPWLLDHRTCPMCKL   DVIKALGYWGEPGDVQEMPAPESPPGRDPAA   NLSLALPDDGSDESSPFSASFASEPQCDPSF   KGDAGENTALLEAGRSDSRHGGPIS   KGDAGENTALLEAGRSDSRHGGPIS   GKSELLKSGSKSTLKHIWTESSKDLSISRILS   QTFRGKENDTDLDLRYDTPEPYSEQDLWDW   LRNSTDLQEPRPRAKRPIVKTGKFKKMFGW   GDPHSNIKTVALLTGRIVDHGNGTSVYF   RHNSTGGGNVSVSLVPPTKIVEFDLAQQTVID   AKDSKSFNCRIEYEKVDATKNTLCNYDPSK   TCYQEQTQSHVSWLCSKPFKVICTYISFYSTD   YKLVQKVCPDYNYHSDTPYFPSG   GGSRGRSDRGSQGDSLYPVGYLDKQVPDTS   VQETDRILVERRCWDLAGFLKQPMNLFIMY   MAGNTISTPTIMMVCMMAWRPIQALMAISAT   FKMLESSSQKFLQGLVYLIGNLMGLALAVYK   CQSMGLLFTHASDWLAFIEPTERMEFSGGGL   LLL   LLL   LLL   LLL   LLL   LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL   SYSVIFIVGLVGNITHYTCHYDEKLLSTVLITT   SYSVIFIVGLVGNITHYTCHYDEKLLSTVLITT   SYSVIFIVGLVGNITHYTCHYDEKLLSTVLITT   SYSVIFIVGLVGNITHYTVTCHYDEKLLSTVLITT   SYSVIFIVGLVGNITHATYTVTCHYDEKLLSTVLITT   SYSVIFIVGLVGNITHALYYVLGJFRKKNSIGYTL   LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL   CKVVGTLPFRIMTHT   CKVVGTLPFRIMTHT   CKVVGTLPFRIMTHT   CKVVGTLPFRIMTHT   CKVVGTLPFRIMTHT   CKVVGTLPFRIMTHT   CKVVGTLPFRIMTHT   CKVVGTLPFR	876	2226	Α	7544	403	587	
877   2227   A   7566   2   940   GCAPDTRFFVPEFGGRGAAPWVALVARGGC							
TFKDKVLVAARRNASAVVLYNEERYGNITLP MSHAGTGNIVVIMISYPKGRELELVQKGIPV TMTGVGTRITVQEFISGQSVVFVAIATIMMII SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI GQLLLHTVKHGEKGIDVDAEBCAVCHENFKV KDIRRIPCKHIPHRICIDPWLDHRTCPMCKL DVIKALGYWGEPGDVQEMPAPESPPGRDPAA NLSLALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS  878  2228  A 7586  315  1232  ERSILCKVDVRWITYSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLTIGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEPDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICTYISFYSTD YKLVQKVCPDYNYHSDTPYPFSG  879  2229  A 7605  479  391  TESWKLKWWSPTCLDQLNGSAPGNVFHG GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGFLKQIPMILFIMY MAGNTISIFPIMMVCMMAWRPIQALMAISAT PKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881  2231  A 7615  291  1452  SPOKTMRSHTIMTTTSVSSWPYSSERRMRFIT NHSDQPPQNFSATPNVITCPMDEKLLSTVLT SYSVIFIVGLVGNIMALYVFLGHKKRNSIQYTL LNVAIADLLLIFCLPFRIMYHNONKWTLGVIL CKVVGTLFRIMYHNNONKWTLGVIL LNVAIADLLLIFGLFRIMYHNONKWTLGVIL LNVAIADLLLIFGLFRIMYHNONKWTLGVIL LNVAIADLLLIFGLFRIMYHNONKWTLGVIL CKVVGTLFFRIMYHNINNIKLIRSI QQRKATTKQSIYVCCLVWMLALGGFLTMIL	877	2227	Α	7566	2	940	
MSHAGTGNIVVIMISYPKGREILELVQKGIPY TMTIGVGTRHVQEFISQGSVSTKALTKKUTKYU GQLLLHTVKHGEKGIDVDAENCAVCIENFKV KDIRILPCKHIFHRICIDPWLLDHRTCPMCKL DVIKALGYWGEPGDVQEMPAPESPPGRDPAA NLSLALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS  878 2228 A 7586 315 1232 ERSILLCKVDVRWTYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHHWTESSKDLSISRILS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRPIVKTGKFKKMPGW GDFHSNIKTVKLINLLITGKIVDHGNGTFSVVF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFEVICTYISFYSTD YKLVQKVCPDYNYHSDTPYPFSG  879 2229 A 7605 479 391 TESWKLKWSPTCLDQLNGSAPGNVFHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQPMILFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT PKMLESSSQKFLQGLVYLIGNLMGLALAYYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTSVSSWYSSERMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQYL LNVAIADLLLIFCLFPRIMYHNQNKWTLGVIL CKVVGTLFRIMYHNMYISIILLGFISLDRYIKINRSI QQRKATTKQSIYVCCIVWMLALGGFLTMIIL							
TMTIGVGTRHVQEFISGQSVYPVAIAFITIMMII SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI GQLLLHTVKHGEKGGDVDAENCAVCIENFKV KDIIRILPCKHIFHRICIDPWLLDHRTCPMCKL DVIKALGYWGEPGDVQEMPAPESPPGRDPAA NLSLALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS  878 2228 A 7586 315 1232 ERSILCKVDVRWIYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTI.KHIWTESSKDLSISRILS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLINLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICTYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG  879 2229 A 7605 479 391 TESWKLKWWSFTCLDQLNGSAPGNVFHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANGGRFKWALELSGPG GGSRGRSDRGSQQOBSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT PKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLIPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFII NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVETVGLVGNIIALYVFLGHRKRNSIQTYL LNVAIADLLLIFCLPFRRMYHNQNKWTIGVIL CKVVGTLFYMNMYISIILGFISLDRYIKINRSI							
GQLLLHTVKHGEKGIDVDAENCAVCIENFKV   KDIRILPCKHIPHRICIDPWLLDHRTCPMCKL   DVIKAALGYWGEPGDVQEMPAPESPPGRDPAA   NLSLALPDDDGSDESSPPSASPAESEPQCDPSF   KGDAGENTALLEAGRSDSRHGGPIS   KGDAGENTALLEAGRSDSRHGGPIS   ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR   RGRMQAACWYVLFLLQFTVYLVTCANLTNG   GKSELLKSGSSKSTLKHIWTESSKDLSISILLS   QTFRGKENDTDLDLRYDTPEPYSEQDLWDW   LRNSTDLQENPRPAKRPIVKTGKFKKMFGW   GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF   RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID   AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK   TCYQEQTQSHVSWLCSKPFKVICTYISFYSTD   YKLVQKVCPDYNYHSDTPYFPSG   S79   2229   A 7605   479   391   TESWKLKWWSPTCLDQLNGSAPGNVFHG   GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS   VQETDRILVEKRCWDIALGPLKQIPMNLFIMY   MAGNTISIFPTIMMVCMMAWRPIQALMAISAT   FKMLESSSQKFLQGLVYLIGNLMGLALAVYK   CQSMGLLPTHASDWLAFIEPPERMEFSGGGL   LL     881   2231   A 7615   291   1452   SPQKTMRSHTITMTTTSVSSWPYSSHRMRFFIT   NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT   SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL   LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL   LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL   LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL   LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL   CKVVGTLFFRIMYHINGNIKWTLGVIL   CKVVGTLFFRIMYHINGNIKWTLGVIL   CKVVGTLFFRIMYHINGNIKWTLGVIL   CKVVGTLFFRIMYHINGNIKWTLGVIL   CKVVGTLFFRIMYHINGNIKWTLGVIL   CKVVGTLFFRIMYHINGNIKWTLGVIL   CKVVGTLFFRIMYHINGNIKWTLGVIL   CKVVGTLFFRIMYHINGNIK TITCHT   CKVVGTLFFRIM THE TATT   CKVVGTLFFRIM THE TATT   CKVVGTLFFRIM THE TATT   CKVVGTLFFRIM THE TATT   CKVVGTLFFRIM THE TATT   CKVVGTLFFRIM THE TATT   CKVVGTLFFRIM THE TATT							
KDIRILPCKHIFHRICIDPWLLDHRTCPMCKL   DVIKALGYWGEPGDVQEMPAPESPPGDPAA   NLSLALPDDOGSDESSPPSASPAESEPQCDPSF   KGDAGENTALLEAGRSDSRHGGPIS   KGDAGENTALLEAGRSDSRHGGPIS   KGDAGENTALLEAGRSDSRHGGPIS   KGDAGENTALLEAGRSDSRHGGPIS   KGDAGENTALLEAGRSDSRHGGPIS   RGRMQAACWYVLFLLQPTVYLVTCANLTING   GKSELLKSGSSKSTLKHIWTESSKDLSISRILS   QTFRGKENDTDLDLR YDTPEPYSEQDLWDW   LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW   GDFHSNIKTVKLINLLITIGKIVDHGNGTFSVYF   RHNDTGGNVSVSLVPPTKIVEFDLAQQTIVID   AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK   TCYQEQTQSHVSWLCSKPFKVICTVISFYSTD   YKLVQKVCPDYNYHSDTYFPPS   YKLVQKVCPDYNYHSDTYFPPS   TESWKLKWWSPTCLDQLNGSAPGNVFIHG   GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS   VQETDRILVEKRCWDIALGFLKQIPMNLFIMY   MAGNTISIFPTIMVCMMAWRPIQALMAISAT   FKMLESSSQKFLQGLVYLIGNLMGLALAVYK   CQSMGLLPTHASDWLAFIEPPERMEFSGGGL   LL     881   2231   A 7615   291   1452   SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT   NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT   SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL   LNVAIADILLIFCLPFRIMYHINQNKWTIGVIL   LNVAIADILLIFCLPFRIMYHINQNKWTIGVIL   LNVAIADILLIFCLPFRIMYHINQNKWTIGVIL   LNVAIADILLIFCLPFRIMYHINQNKWTIGVIL   LNVAIADILLIFCLPFRIMYHINQNKWTIGVIL   LNVAIADILLIFCLPFRIMYHINQNKWTIGVIL   CKVVGTLTKQSIVVCCIVWMLALGGFLTMIIL							
878 2228 A 7586 315 1232 ERSILCKVDVRWIYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFILQPTVYLVTCANLTNG GKSELLKSGSSKSTILKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPPPYSEQDL WDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHISNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPIKIVEFDLAQQTIVD AKDSKSFNCRIEYEKVDKATKNTILCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFFSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSQQDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQFPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNILALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRRMYHNQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITITQSIYVCCIVWMLALGGFLTMIIL							GQLLLHTVKHGEKGIDVDAENCAVCIENFKV
878 2228 A 7586 315 1232 ERSILCKVDVRWIYVSEGIKTQRRHRQGSLR RGRMQAACWYVLFILQFIVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSYVF RHNSTGQGNVSVSLVPPIKIVEFDLAQQIVID AKDSKSPNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKILSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLFCLPFRIMYHINQNKWTLGVIL CKVVGTLFFRIMYHINQNKWTLGVIL CKVVGTLFFRIMYHINQNKWTLGVIL CKVVGTLFFRIMYHINQNKWTLGVIL CKVVGTLFFRIMYHINQNKWTLGVIL OQRKATITKQSIYVCCIVWMLALGGFLTMIIL						1	KDIIRILPCKHIFHRICIDPWLLDHRTCPMCKL
878 2228 A 7586 315 1232 ERSILCKVDVRWIYVSEGIKTQRRHRQGSLR RGRMQAACWYVLFILQFIVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSYVF RHNSTGQGNVSVSLVPPIKIVEFDLAQQIVID AKDSKSPNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKILSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLFCLPFRIMYHINQNKWTLGVIL CKVVGTLFFRIMYHINQNKWTLGVIL CKVVGTLFFRIMYHINQNKWTLGVIL CKVVGTLFFRIMYHINQNKWTLGVIL CKVVGTLFFRIMYHINQNKWTLGVIL OQRKATITKQSIYVCCIVWMLALGGFLTMIIL							DVIKALGYWGEPGDVQEMPAPESPPGRDPAA
878 2228 A 7586 315 1232 ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFILQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRILS QTFRGKENDTDI.DLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICTYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFTHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSRGSQQDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGFLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPFERMEFSGGGL LL SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVITCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							NLSLALPDDDGSDESSPPSASPAESEPQCDPSF
RGRMQAACWYVLFLLQFTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVIID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDITYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSQQDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGFLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFII NHSDQPPQNFSATPNVITCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADILLIFCLPFRIMYHNQNKWTLGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							KGDAGENTALLEAGRSDSRHGGPIS
GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFVGLVGNIIALVYLGJIHRKNSIGIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	878	2228	Α	7586	315	1232	ERSLLCKVDVRWIYVSEGIKTQRRHRQGSLR
QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICTYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFTHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVITCPMDEKLLSTVLTT SYSVIFIVGLVGNILALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWILGVIL CKVVGTLFYMNMYISIILJGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSQ 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT PKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS
GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFTHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRPKWAIELSGPG GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT PKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
RHNSTGQGNVSVSLVPPIKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICTYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG  879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFTHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPILKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT PKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADILLIFCLPFRIMYHINQNKWTLGVIL CKVVGTILFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL						İ	LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW
879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAILSGPG GGSRGSDRGSQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPIMMVCMMAWRPIQALMAISAT PKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLYGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWILGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVITCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWILGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQYL LNVAIADLLLIFCLPFRIMYHINQNKWILGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL						•	
880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT PKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTTTMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	2.50						
GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGFIK,QIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT PKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	880	2230	Α	7612	93	659	DAAVAMTAQGGLVANRGRRFKWAIELSGPG
VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS
881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL		-		.			VQETDRILVEKRCWDIALGPLKQIPMNLFIMY
CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVITCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWILGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							MAGNTISIFPTMMVCMMAWRPIQALMAISAT
881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVITCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIL							
881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVITCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL						İ	
NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							LL
NHSDQPPQNFSATPNVITCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWILGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	881	2231	Α	7615	291	1452	SPOKTMRSHTITMTTTSVSSWPYSSHRMRFIT
SYSVIFIVGLYGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL				7			SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL
CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL						·	
QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL					ļ	·	
TI KKGGHNSTMCFHVRDKHNAKGFATENTI					ļ		QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL
							TLKKGGHNSTMCFHYRDKHNAKGEAIFNFIL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	DNO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	.nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		(		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
		1				VVMFWLIFLLIILSYIKIGKNLLRISKRRSKFPN
1 .						SGKYATTARNSFIVLJIFTICFVPYHAFRFTYISS
1						QLNVSSCYWKEIVHKTNEIMLVLSSFNSCLDP
						VMYFLMSSNIRKIMCQLLFRRFQGEPSRSEST
882	2232	A	7617	67	379	SEFKPGYSLHDTSVAVKIQSSSKST
802	22,32	^	/01/	67	3/9	RQMALLKANKDLISAGLKEFSVLLNQQVFND
1				•		PLVSEEDMVTVVEDWMNFYINYYRQQVTGE
						PQERDKALQELRQELNTLANPFLAKYRDFLK SHELPSHPPPSS
883	2233	A	7622	400	215	KVKTCRYNPKYSAANDTGFVDIPSREKDLAK
502	2233	1.	1022	100	213	AVATVGPISVAVGASHVFFQFYKKGKHLSS
884	2234	Α	7638	2640	2861	APVLILQMVKLSIVLTPQFLSHDQGQLTKELQ
		. **	7000	2010	,2001	QHVKSVTCPCEYLRKVSECROMGPGALEOFP
ļ.						GLSCHTSHSG
885	2235	Α	7642	201	455	PSRGKMELEAMSRYTSPVNPAVFPHLTVVLL
			70.2		133	AIGMFFTAWFFVYEVTSTKYTRDIYKELLISL
						VASLFMGFGVLFLLLWVGIYV
886	2236	A	7692	61	569	APENPPSRQHFNSETKVKLSLKTGTWLGNHA
1					1	HLGEHFSTHHELGLSGKVVGFLVKNILEVIRN
						GGMETRHPGKVSSWFHRWDSRAEQHNHAE
]					-	HHEDVPQGDEDSKVSEAQQEFPDVVTCAGLP
]						GLLPKALRVLLFQLKVQHRPGIHQQRPEQQD
						VSDHRYGRSVRQNRK
887	2237	Α	7693	85	315	NPGCCLPVAMRTSYLLLFTLCLLLSEMASGG
1 1						NFLTGLGHRSDHYNCVSSGGQCLYSACPIFTK
l						IQGTCYRGKAKCCK
.888	2238	A	7702	242	1298	APSHRRRYLSPSRSAGQLGNMALERLCSVLK
1 1						VLLITVLVVEGIAVAQKTQDGQNIGIKHIPAT
<b>!</b>						QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI
						LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR
						DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF
[ [				·		SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP
]						DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF
1 1						VAVYDGSSSIENLKAKFCSTVANDVMLKTGI
		1		ļ		GVIRMWADEGSRLNRFRMLFTSFGGASPAOA
1			- 1		1	ALSFCHSNMCINNSLVCNGVQNCAYPWDEN
						HC
889	2239	Α	7707	185	2911	CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE
		1			İ	MTMDALLARLKLLNPDDLREEIVKAGLKCGP
	- 1	1	1	ł		ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA
[ l		1				GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF
1		l				GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG
, ,		1	j	1	. [	ATASKEPPLYYGVCPVYEDVPARNERIYVYE
		l				NKKEALQAVKMIKGSRFKAFSTREDAEKFAR
	· [					GICDYFPSPSKTSLPLSPYKTAPLFSNDRLKDG
		l	,			LCLSESETVNKERANSYKNPRTQDLTAKLRK
	1	1	Í	}		AVEKGEEDTFSDLIWSNPRYLIGSGDNPTTVQ
		i				EGCRYNVMHVAAKENQASICQLTLDVLENP
	Ţ	l				DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP
	1	Į	J		. ]	DKMGYDTPLHFACKFGNADVVNVLSSHHLI
	į	- 1	ļ			VKNSRNKYDKTPEDVICERSKNKSVELKERIR
		j	ļ			EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA
		Ì				EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE
ľĺ	ĺ	ĺ	ĺ		ľ	DFRKLWKTPPREKAGFLHHVKKSDPERGFER
	[		Į	ļ	į	VGRELAHELGYPWVEYWEFLGCFVDLSSQE
	1		l	1		GLQRLEEYLTQQEIGKKAQQETGEREASCRD
		l	i			KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA
LL	لـــــــا					ARNNSPPTVGAFGHTRCSAFPLEQEADLIEAA

NO: of nucl- cotide seq-	SEQ ID NO: of peptide	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl- cotide seq-		DOG	ID NO:	heamnna		
cotide seq-	pepuae		•			D=Aspartic Acid, E=Ghrtamic Acid,
seq-			in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
- 1	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	ncuce		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
}	- 1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i				peptide		/=possible nucleotide deletion, \=possible
1				Sequence		nucleotide insertion
						EPGGPHSSRNGLCHPLNHSRTLAGKRPKAPR
					}	GEEAHLPPVSDLTVEFDKLNLQNIGRSVSKTP
1	ļ					DESTKTKDOILTSRINAVERDLLEPSPADOLG
j	j					
	1					NGHRRTESEMSARIAKMSLSPSSPRHEDQLEV
						TREPARRLFLFGEEPSKLDQDVLAALECADV
i						DPHQFPAVHRWKSAVLCYSPSDRQSWPSPAV
·	i					KGRFKSQLPDLSGPHSYSPGRNSVAGSNPAKP
						GLGSPGRYSPVHGSQLRRMARLAELAAL
890	2240	Α	7711	360	269	RHMPVIPALWEAEVGGLLEPRSSRSAWATE
891	2241	Α	7721	61	1175	KLPWEPSFLIKMQIIRHSEQTLKTALISKNPVL
				-		VSQYEKLDAGEQRLMNEAFQPASDLFGPITL
- 1	i	i				HSPSDWITSHPEAPQDFEQFFSDPYRKTPSPN
- 1	į	1		٠.		KRSTYIQSIGSLGNTRIISEEYIKWLTGYCKAYF
ŀ	1					YGLRVKLLEPVPVSVTRCSFRVNENTHNLQIH
i	ł	1	- 1			
1	ļ	ļ		•		AGDILKFLKKKKPEDAFCVVGITMIDLYPRDS
ŧ	l l					WNFVFGQASLTDGVGIFSFARYGSDFYSMHY
J	j	j	J		•	KGKVKKLKKTSSSDYSIFDNYYIPEITSVLLLR
	1	l				SCKTLTHEIGHIFGLRHCQWLACLMQGSNHL
	į.					EEADRRPLNLCPICLHKLQCAVGFSIVERYKA
į	l l		- 1			LVRWIDDESSDTPGATPEHSHEDNGNLPKPV
						EAFKEWKEWIIKCLAVLQK -
892	2242	Α	7723	2	1650	SAPTAPARPCRAERGSGGGMLALLAASVALA
- 1						VAAGAQDSPAPGSRFVCTALPPEAVHAGCPL
- 1	- 1	ł	- 1			PAMPMQGGAQSPEEELRAAVLOLRETVVQQ
1	j	1				KETLASARAIRELTGKLARCEGLAGGKARGA
	i	[				GATGKDTMGDLPRDPGHVVEQLSRSLQTLK
ł	- 1	l	1			DRLESLEPLPAMPMQGGAQSPEEELRAAVLQ
	ŀ	İ	İ			LRETVVQQKETLASARAIRELTGKLARCEGL
			- 1			
į.		ŀ	J	J		AGGKARGAGATGKDTMGDLPRDPGHVVEQ
1	- 1					LSRSLQTLKDRLESLEHQLRANVSNAGLPGD
į		i				FREVLQQRLGELERQLLRKGAELEDEKSLLH
	- 1	i				NETSAHRQKTESTLNALLQRVTELERGNSAF
ĺ	T I	- (		1		KSPNAFKVSLPLRTNYLYGKIKKTLPELYAFT
		- 1				ICLWLRSSASPGMGTPFSYAVPGQANEIVLIE
	ļ	i	ľ			WGNNPIELLINDKVAQLPLFVSDGKWHHICV
į.	- 1	- 1	İ	ł	•	TWTTRDGMWEAFQDGKKLGTGENLAPWHPI
1		-	l			KPGGVLILGQEQDTVGGRFDATQAFVGELSO
1	i	1	ľ	ľ		FNIWDRVLRAQEIVNIANCSTNMPGNIIPWVD
		- 1	l	· · · · · · · · · · · · · · · · · · ·		NNVDVFGGASKWPVETCEERLLDL
893 2	2243	A	7729	3554	2419	LTAGTAMNYPLTLEMDLENLEDLFWELDRL
-						DNYNDTSLVENHLCPATEGPLMASFKAVFVP
	.1	- 1	j	. <b>!</b>		
	-		l			VAYSLIFLLGVIGNVLVLVILERHRQTRSSTET
1	]	Į	J	. ]	J	FLFHLAVADLLLVFILPFAVAEGSVGWVLGTF
J	. 1			l		LCKTVIALHKVNFYCSSLLLACIAVDRYLAIV
1			ì	į	İ	HAVHAYRHRRLLSIHITCGTIWLVGFLLALPEI
i	1	1	J	ļ		LFAKVSQGHHNNSLPRCTFSQENQAETHAWF
ļ	İ	- 1		1		TSRFLYHVAGFLLPMLVMGWCYVGVVHRLR
1	}	ł	ļ	ļ		QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV
j	ļ		1	İ		IFLDTLARLKAVDNTCKLNGSLPVATTMCEFL
1	ł	l	ł	i	1	GLAHCCLNPMLYTFAGVKFRSDLSRLLTKLG
ľ						CTGPASLCQLFPSWRRSSLSESENATSLTTF
894 2	2244	A	7738	670	287	FVTRAGRWGAGARVRGGAGGMASGAARWL
·   1			.,,,,,	١,٠٠	201	
1	l	1		l		VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR
1	l	ŀ	j	l		SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI
.	}	j	]	j	ļ	VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD
						D
895 2	2245	Α	7753	119	278	APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR
			ļ			LWLSLFLHAGKEAPHCPRTRPL
						SPAWWNSQQRVVSPFLALLTLEPTFHHLLPIM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isolcucine, K=Lysine, L=Lcucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  QVSTAALAVILCTMALCNQVLSAPLAADTPT ACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFL
897	2247	A	7761	1725	445	TKRGRQVCADPSEEWVQKYVSDLELSA RPRRRGTHHFSCVLGSFRVSAMFPRVSTFLPL RPLSRHPLSSGSPETSAAAMILLTVRHGTVRY RSSALLARTKNNIQRYFGTNSVICSKKDKQSV RTEETSKETSESQDSEKENTKKDLLGIIKGMK VELSTVNVRTTKPPKRRPLKSLEATLGRLRRA TEYAPKKRIEPLSPELVAAASAVADSLPFDKQ TTKSELLSQLQQHEEESRAQRDAKRPKISFSNI ISDMKVARSATARVRSRPELRIQFDEGYDNYP GQEKTDDLKKRKNIFTGKRLNIFDMMAVTKE APETDTSPSLWDVEFAKQLATVNEQPLQNGF EELIQWTKEGKLWEFPINNEAGFDDDGSEFH EHIFLEKHLESFPKQGPIRHFMELVTCGLSKNP YLSVKQKVEHIEWFRNYFNEKKDILKESNIQF KLRPWKFLFRNN
898	2248	A	7775	85	496	SCQTTQPPAQSCSTGTMRIMLLFTAILAFSLA QSFGAVCKEPQEEVVPGGGRSKRDPDLYQLL QRLFKSHSSLEGLLKALSQASTDPKESTSPEK RDMHDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKYPPRAE
899	2249	A	7785	179	703	PFHLGASSNTFRLQVQTQESKAQKEVKMGFI FSKSMNESMKNQKEFMLMNARLQLERQLIM QSEMRERQMAMQIAWSREFLKYFGTFFGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQYDL GYGTLLERMKGEAEDILETEKSKLQLPRGMIT FESIEKARKEQSRFFIDK
900	2250	A	7789	1465	300	VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNPSSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN
901	2251	A	7796		807	VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR
902	2252	<b>A</b>	7802	2	721	TAARROKGTAARRLOKGTAARRROKGTAA RRROKGTAARRPOKGTAARRROKGTAARRR QKGTAARRPOKGTAARRROKG TAARRROKGTAARRPOKGTAARRROKG TAARROKGTAARRROKGLAIASRGCPCASR AGGVRGAGSRLRAMAPKVFROYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG AAACVYFGIAASLVKMGRLEGWEVFAKPKV

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
į		ł		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide		/=possible nucleotide deletion, \-possible
				sequence		nucleotide insertion
903	22.53	A	7807	1	584	PWLPWSDGRAARSSRKCPRSRFPVQVGKMA
1						VSTVFSTSSLMLALSRHSLLSPLLSVTSFRRFY
1					ľ	RGDSPTDSQKDMIEIPLPPWQERTDESIETKR
1						ARLLYESRKRGMLENCILLSLFAKEHLQHMT
1		ļ				EKQLNLYDRLINEPSNDWDIYYWATEAKPAP
i i						EIFENEVMALLRDFAKNKNKEQRLRAPDLEY
904	2254	A	7813	40	821	LFEKPR CACDA COVA A A DADA COMPANIA
1	22.74	Λ	7013	40	021	GAGRALGHLETGAGDVAAALPARKFPRSLLG
						AGARLTGWTMNVFRILGDLSHLLAMILLLGK
		1				IWRSKCCKGISGKSQILFALVFITRYLDLFTNF
i 1						ISIYNTVMKVVFLLCAYVTVYMIYGKFRKTF DSENDTFRLEFLLVPVIGLSFLENYSFTLLEIL
						WTFSIYLESVAILPOLFMISKTGEAETITTHYL
						FFLGLYRALYLANWIRRYQTENFYDQIAVVS
1. (	·					GVVQTIFYCDFFYLYVTKGRSWDDSNADTGI.
						RSYSSI
905	2255	Α	7817	1399	881	LSNKDVLSPQLKDENSKLRRKLNEVOSFSEA
						QTEMVRTLERKLEAKMIKEESDYHDLESVVO
			•			QVEQNLELMTKRAVKAENHVVKLKQEISLL
{ {						QAQVSNFQRENEALRCGQGASLTVVKQNAD
1 1						VALQNLRVVMNSAQASIEQLVSGAETLNLVA
						EILKSIDRISEVKDEEEDS
906	2256	Α	7822	3	1462	DSPRNRFEILGRPTRTPTRPGPRPAMEDLDAL
ļ						LSDLETTTSHMPRSGAPKERPAEPLTPPPSYG
!						HQPQTGSGESSGASGDKDHLYSTVCKPRSPK
1						PAAPAAPPFSSSSGVLGTGLCELDRLLQELNA
]						TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK
l i			i			RPSLPSSPSPGLPKASATSATLELDRLMASLSD
! !				l		PRVQNHLPASGPTQPPVVSSTNEGSPSPPEPTG
				j		KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC
[ [	ĺ	' i	- 1		İ	NKPIAGQVVTALGRAWHPEHFVCGGCSTAL
l I						GGSSFFEKDGAPFCPECYFERFSPRCGFCNQP1
					ļ	RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN
1 1	1		}		İ	YISALSALWHPDCFVCRECFAPFSGGSFFEHE
	i					GRPLCENHFHARRGSLCATCGLPVTGRCVSA
	1					LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY
	j					CQPCFLKLFG
907	2257	Α	7828	1792	1671	FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH
	ĺ			-	,	YCKSQAWG
908	2258	A	7842	110	1172	KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG
[	ł					SLOPPPSGLKOSSHLSLSSSWDFRHAPTHPET
	İ	ľ	i			YTCPKMIEMEQAEAQLAELDLLASMFPGENE
		- 1	1			LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI
			l			NMNLDVSDEKMAMFSLACILPFKYPAVLPEI
	ļ	· [	j		•	TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV
l İ	i		j	•		CILNATEWVREHASGYVSRDTSSSPTTGSTVQ
	ļ	[	. [	ĺ	ĺ	SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL
'					ļ	SLSGFSMPGKPGVVCVEGPQSACEEFWARLR
						KLNWKRILIRHREDIPFDGTNDETERQRKFSIF
ŀ	ļ	ł	ł	1	ł	<b>EEKVFSVNGARGNHMDFGQLYQFLNTKGCG</b>
<u> </u>		I				DVFQMFLWV
909	2259	A	7870	3067	2923	EGICVYTFIYVHMYTRTCMHTYPYMYMNSV
			100			LISSEILLIPSKYLFESK
910	2260	Α	7884	212	4874	GALTWSHPLLAVCPQGVWLGSTPSGSPALLP
	i	- 1	- 1	ľ		PSHRVNAEPGCVVTNACASGPCPPHANCRDL
		l		l	1	WQTFSCTCQPGYYGPGCVDACLLNPCQNQG
		!	ļ	l	İ	SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD
		l	l			QQCPRGWWGSPTCGPCNCDVHKGFDPNCNK

NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide peptide sequence  NO: of peptide sequence  No: of peptide	z, H=Histidine, Leucine, ne, P=Proline, S=Serine, =Tryptophan, =Stop codon,
ectide sequence USSN location corresponding to last amino acid residue of peptide sequence sequence sequence sequence over the sequence sequence sequence over the sequence sequence sequence sequence sequence corresponding to last amino acid residue of peptide sequence sequence sequence sequence sequence sequence corresponding to last amino acid residue of peptide sequence sequen	.eucine, ne, P=Proline, S=Serine, =Tryptophan, =Stop codon,
uence 914 ng to first amino acid residue of peptide residue of peptide residue of peptide sequence sequence 914 ng to first amino acid residue of peptide residue of peptide sequence y=Tyrosine, X=Unknown, * /=possible nucleotide deletion nucleotide insertion	S=Serine, =Tryptophan, =Stop codon,
amino acid residue of peptide residue of peptide sequence T=Threonine, V=Valine, W= Y=Tyrosine, X=Unknown, *  peptide sequence sequence nucleotide insertion	Tryptophan, Stop codon,
residue of peptide sequence Y=Tyrosine, X=Unknown, *  /=possible nucleotide deletion nucleotide insertion	=Stop codon,
peptide /=possible nucleotide deletion sequence nucleotide insertion	
sequence nucleotide insertion	n, ≔possible
TNGQCHCKEFHYRPRGSI	OCCUPATION DISCOURSE
SRSCAPHSGQCPCRPGAL	GROCNSCOSPEARY
TASGCRVLYDACPKSLRS	GVWWPOTKFGVI.
ATVPCPRGALGLRGAGA	
PDLFNCTSPAFRELSLLLD	
AKKLAQRLREVTGHTDH	
AHLLAFESHQQGFGLTAT	
GSALLAPETGDLWAALG	-
RHLEEYAATLARNMELT	
SIDRMEHPSSPRGARRYPI	
DPHTHVLLPSQSPRPSPSE	VLPISSSIENSTISS
VVPPPAPPEPEPGISIIILLV AERRGARLPONPVMNSPV	
GILESPISLEFRLLQTANRS	
EQHGVWTARDCELVHRN	IGSHARCRCSRTGT
PGVLMDASPRERLEGDLE	
VAALVLTAAILLSLRSLKS	
LGVAELLFLLGIHRTHNQI	LVCTAVVILLHYFF
LSTFAWLFVQGLHLYRM	
FYHALGWGVPAVLLGLA	
CWISVHEPLIWSFAGPVVI	
ARTSCSTGQREAKKTSAL	TLRSSFLLLLLVSA
SWLFGLLAVNHSILAFHY LLLFCVLNADARAAWMP	
PAPGLGPGAYNNTALFEE	
VSSARSGRTQDQDSQRGR	RSYLRDNVLVRHGS
AADHTDHSLQAHAGPIDI	
DSDSDSDLSLEEERSLSIPS	
QRPLCRAAQSERLLTHPK	
LGECEAAPCALQTWGSER	
NNQPDPALTSGDETSLGR	
QYPLVPQTRGAPELSWCR	
YGRIYAGGGTGSLSQPASI	`
QLSRERLEEAPAPVLRPLS	•
RLEPKDRGSTLPRRQPPRI DALDLGAPREWLSTLPPPI	
LSPQRQLSRDPLLPSRPLD	
VPSRHPSREALGPLPQLLR	
STEQLDILSSILASFNSSAL	SSVQSSSTPLGPHT
TATPSATASVLGPSTPRSA	TSHSISELSPDSEPR
DTQALLSATQAMDLRRRI	DYHMERPLLNQEH
LEELGRWGSAPRTHQWR	
LLQHLPVLVWLPRYPVRD	
IMQLPQGLAYALLAGLPPI	
911 2261 A 7890 21 806 EFGTSRSSRSMAEDLGLSF	
SCRPKARSSSARWALTCC	
LVSQLRAQGEACVOFOAL	
YAPLRADGDKPRAHLTVV	
ALHWEHELGLAFTKNRMI	
YFIYSQVTFRGMTSECSEI	
VITKVTDSYPEPTQLLMGT	•
PIYLGAMFSLQEGDKLMV	- 1
DKTFFGAFLL	1
912 2262 A 7891 1263 111 ACGIRHEGALPGLTATPEA	MLRFLPDLAFSFL
LILALGQAVQFQEYVFLQF	
PVPYILKKIFQDREAAATT	
VRGNVLRFLPDQGFFLYPI	
YFNLSAIKEREQLTLAQLG	
ELELALFLVQEPHVWGQT	ILVLOVWILATK2A

SEQ ID	SEQID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		USSN	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence	ĺ	09/496	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
uence	uaice		914	ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
l donce		l	714	amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine,
1				residue of	sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
J		J		peptide	sequence	/=possible nucleotide deletion, \=possible
:		ļ		sequence		nucleotide insertion
		<del> </del>	<del>                                     </del>	Sequence	<del></del>	PWPQGAVHFNLLDVAKDWNDNPRKNFGLFL
l						EILVKEDRDSGVNFQPEDTCARLRCSLHASLL
	1			l		VVTLNPDQCHPSRKRRAAIPVPKLSCKNLCH
	İ			i		RHQLFINFRDLGWHKWIIAPKGFMANYCHGE
1		ļ	i		1	CPFSLTISLNSSNYAFMQALMHAVDPEIPQAV
j	}	İ	]	i	]	CIPTKLSPISMLYQDNNDNVILRHYEDMVVD
	·	Ì			ļ	ECGCG
913	2263	Α	7892	15	849	ASRLPRGPGCGADMRPLLGLLLVFAGCTFAL
						YLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAE
i						LRELSEVLREYRKEHQAYVFLLFCGAYLYKQ
						GFAIPGSSFLNVLAGALFGPWLGLLLCCVLTS
		[	İ			VGATCCYLLSSIFGKQLVVSYFPDKVALLOR
1					1	KVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPI
T I			1			LNIPIVQFFFSVLIGLIPYNFICVQTGSILSTLTS
	ŀ					LDALFSWDTVFKLLAIAMVALIPGTLIKKFSQ
						KHLQLNETSTANHIHSRKDT
914	2264	A	7893	815	959	KSGWVWWLTPLIPALWEAQTEGSLRPEVKN
015	2000					RLSNITRPFFSKKKKILV
915	2265	A	7909	3	641	HASGPGGLLRRRRGSGANMPVARSWVCRKT
						YVTPRRPFEKSRLDQELKLIGEYGLRNKREV
J						WRVKFTLAKIRKAARELLTLDEKDPRRLFEG
1						NALLRRLVRIGVLDEGKMKLDYILGLKIEDFL
						ERRLQTQVFKLGLAKSIHHAHVLIQQCHIRVR
						EQVVNILFFTVRLDSQKHIDFSLCFPIGVANPS
916	2266	A	7914	3	967	HVKRKNASKGQGGAGARDDEEEE VAHTQWHTCQRLSQLTHRSILKYLLIDTHAC
			// /	3	507	QVLILKHTHASLSLPSCQECFPSSIPSASHMVS
i						HPHPPPSPRWGQTPEGLPAASPCGPGPRSCFS
1 1						SILPTGDSWGMLACLCTVLWHLPAVPALNRT
						GDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLN
						YLGPPFNEPDFNPPRLGAETLPRATVDLEVW
) )	J		]		j	RSLNDKLRLTQNYEAYSHLLCYLRGLNRQAA
						TAELRRSLAHFCTSLQGLLGSIAGVMAALGY
1		ļ		I		PLPQPLPGTEPTWTPGPAHSDFLQKMDDFWL
		1		i		LKELQTWLWRSAKDFNRLKKKMQPPAAAVT
017	200	<del></del>				LHLGAHGF
917	2267	A	7921	2	1166	RPRRGQGLVQEVQTENVTVAEGGVAEITCRL
	<b> </b>					HQYDGSIVVIQNPARQTLFFNGTRALKDERFQ
	}					LEEFSPRRVRIRLSDARLEDEGGYFCQLYTED
						THHQIATLTVLVAPENPVVEVREQAVEGGEV
}	ĺ	- 1	1	İ	•	ELSCLVPRSRPAATLRWYRDRKELKGVSSSQ
		i	1			ENGKVWSVASTVRFRVDRKDDGGIIICEAQN
	1	ł	l	0.0		QALPSGHSKQTQYVLDVQYSPTARIHASQAV VREGDTLVLTCAVTGNPRPNQIRWNRGNESL
1 1	1		l			PERAEAVGETLTLPGLVSADNGTYTCEASNK
		. 1	1			HGHARALYVLVVYGESRLRPTEGGGGAPDP
1	Ì	I				GAVVEAQTSVPYAIVGGILALLVFLIICVLVG
		1			1	MVWCSVRQKGSYLTHEASGLDEQGEAREAF
	ſ	ľ	ŀ	Ĭ	ŀ	LNGSDGHKRKEEFFI
918	2268	A	7938	3	2653	RRRLPPASPPSSSVSSSLSPSAVVMACRWSTK
				-		ESPRWRSALLLLFLAGVYGNGALAEHSENVH
•	}	. l	l	ł	ļ	ISGVSTACGETPEQIRAPSGIITSPGWPSEYPAK
	}	l	1	l		INCSWFIRANPGEITTISFODFDIOGSRRCNLD
	İ	ļ	ļ		l	WLTIETYKNIESYRACGSTIPPPYISSODHIWIR
1 1		j	]	}	• [	FHSDDNISRKGFRLAYFSGKSEEPNCACDQFR
		1	j	ĺ		CGNGKCIPEAWKCNNMDECGDRSDEEICAKE
				ľ	.	ANPPTAAAFQPCAYNQFQCLSRFTKVYTCLP
				]	l	ESLKCDGNIDCLDLGDEIDCDVPTCGQWLKY
L					l	FYGTFNSPNYPDFYPPGSNCTWLIDTGDHRK
			<del></del> -			

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	Elsoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
vence	i	Ì	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		i	l	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
i l			1	peptide		possible nucleotide deletion, possible
			i	sequence		nucleotide insertion
						VILRETDEKLDGTGYGDYVKIYDGLEENPHK
1 1				ł .	1	LLRVLTAFDSHAPLTVVSSSGOIRVHFCADKV
<b>!</b>			1			NAARGFNATYQVDGFCLPWEIPCGGNWGCY
1 . !		·		]		TEQQRCDGYWHCPNGRDETNCTMCQKEEFP
1						CSRNGVCYPRSDRCNYQNHCPNGSDEKNCFF
ł i				ļ	1	CQPGNFHCKNNRCVFESWVCDSQDDCGDGS
				1		DEENCPVIVPTRVITAAVIGSLICGLLLVIALG
						CTCKLYSLRMFERRSFETQLSRVEAELLRREA
1 1						PPSYGQLIAQGLIPPVEDFPVCSPNQASVLENL
1 1					]	RLAVRSQLGFTSVRLPMAGRSSNIWNRIFNFA
}						RSRHSGSLALVSADGDEVVPSQSTSREPERNH
1 1					1 1	THRSLFSVESDDTDTENERRDMAGASGGVAA
]						PLPQKVPPTTAVEATVGACASSSTOSTRGGH
				İ		ADNGRDVTSVEPPSVSPARHOLTSALSRMTO
1 I					į į	GLRWVRFTLGRSSSLSONOSPLROLDNGVSG
1 1					ĺ	REDDDDVEMLIPISDGSSDFDVNDCSRPLLDL
1 1						ASDQGQGLRQPYNATNPGVRPSNRDGPCERC
			i l			GIVHTAQIPDTCLEVTLKNETSDDEALLLC
919	2269	Α	7951	1674	1839	VVRVTCCPPARSTTERTNAYDEEDCVEMVAS
1 1						GGWNDVACHTTMYFMCEFDKKNM
920	2270	Α	7953	47	572	GGRASWPEQAKEPRREGHTDKQQTEDVLAA
	1					GLRCLPHLPAICARRMSPAFRAMDVEPRAKG
						VLLEPFVHQVGGHSCVLRFNETTLCKPLVPRE
		l	İ			HQFYETLPAEMRKFTPQYKGKSQLLEGLPHW
] ]	j	1				RGDVRDRGHGRPWQPSLEPSLPPTLCFPSLSS
LI		[				FSSSWPSAQHLTPSVFNPW
921	2271	Α	7957	612	812	RSGRTVVTGIGYSKALQSSNRNTKSLLQNEF
i i	į .	- 1				MMVYSFRALSFKESTWATFQHGGEATKSRSL
						SSTQ
922	2272	A	7967	1443	1660	ENITEKWKEIWMCRGNKKSCCWTFIKDRHLT
1 [						VSCCKSKSGETLLICIFCSNLVGFFFFGIRGFSN
						WELVKPN
923	2273	Α	7981	1	3023	GSAPRAATAMARARPPPPPSPPPGLLPLLPPLL
<u> </u>		l				LLPLLLLPAGCRALEETLMDTKWVTSELAWT
t I	[	į			1	SHPESGWEEVSGYDEAMNPIRTYQVCNVRES
	ŀ	- 1	i			SQNNWLRTGFIWRRDVQRVYVELKFTVRDC
	ĺ	f				NSIPNIPGSCKETFNLFYYEADSDVASASSPFW
	İ			·		MENPYVKVDTIAPDESFSRLDAGRVNTKVRS
	ł	ł	l		ŀ	FGPLSKAGFYLAFQDQGACMSLISVRAFYKK
	ļ	l			·	CASTTAGFALFPETLTGAEPTSLVIAPGTCIPN
	İ	Ì			· 1	AVEVSVPLKLYCNGDGEWMVPVGACTCATG
1 1	· •				;	HEPAAKESQCRPCPPGSYKAKQGEGPCLPCPP
!!	1	J	1			NSRTTSPAASICTCHNNFYRADSDSADSACTT
[			l		. 1	VPSPPRGVISNVNETSLILEWSEPRDLGVRDD
	ļ		·			LLYNVICKKCHGAGGASACSRCDDNVEFVPR
.	ļ		. 1			QLGLSEPRVHTSHLLAHTRYTFEVQAVNGVS
t l	1	1	ŀ			GKSPLPPRYAAVNITTNQAAPSEVPTLRLHSS
j l		l	. 1			SGSSLTLSWAPPERPNGVILDYEMKYFEKSEG
	ſ	ĺ	1	1		IASTVTSQMNSVQLDGLRPDARYVVQVRART
				1	, I	VAGYGQYSRPAEFETTSERGSGAQQLQEQLP
]			ł	j	. 1	LIVGSATAGLVFVVAVVVIAIVCLRKQRHGS
	]	·		}	İ	DSEYTEKLQQYIAPGMKVYIDPTTYEDPNEA
ĺ		- 1	- 1	1	i	VREFAKEIDVSCVKIEEVIGAGEFGEVCRGRL
		1	1			KQPGRREVFVAIKTLKVGYTERQRRDFLSEA
		I	Í	ļ	. [	SIMGQFDHPNIIRLEGVVTKSRPVMILTEFME
			J		. [	NCALDSFLRLNDGQFTVIQLVGMLRGIAAGM
1 1			- 1		ļ.	KYLSEMNYVHRDLAARNILVNSNLVCKVSDF
		J	j		j	GLSRFLEDDPSDPTYTSSLGGKIPIRWTAPEAI AYRKFTSASDVWSYGIVMWEVMSYGERPY

SEQ ID NO: of nucl- ectide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Gfycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  WDMSNQDVINAVEQDYRLPPPMDCPTALHQ LMILDCWVRDRNLRPKFSQIVNTLDKLIRNAA SLKVIASAQSGMSQPLLDRTVPDYTTFTTVGD WLDAIKMGRYKESFVSAGFASFDLVAQMTA EDLLRIGVTLAGHQKKILSSIQDMRLQMNQT
924	2274	A	7985	1	503	LPVQV  FRPRTKKATAMYLEHYLDSIENLPCELQRNF QLMRELDQRTEDKKABIDILAAEYISTVKTLS PDQRVERLQKIQNAYSKCKEYSDDKVQLAM QTYEMVDKHIRRLDADLARFEADLKDKMEG SDFESSGGRGLKKGRGQKEKRGSRGRGRRTS EEDTPKKKKHKGG
925	2275	A	7994	447	589	LPCSFCAQCMSSFERVWLQQSHFHNPRWNSR SPIRCYCQHWPHCVHC
926	2276	A	7996	925	582	GPCKVCCITLAIMLQCHSFYRKDVQVEHPKS LNPKYSQIENFLSADMALKRKCLLSISDLDFW IWDAQPVGIMQTILQNLKKIPNPGCFWSQAFQI RDTQPILPLGGRYYITIRO
927	2277	A	7998	2	353	RIQRPLNSRSPNHSLFVKAELTAKQATMKLSV CLLLVTLALCCYQANAEFCPALVSELLDFFFI SEPLFKLSLAKFDAPPEAVAAKLGVKRCTDQ MSLQKRSLIAEVLVKILKKCSV
928	2278	A	8004	130	588	LAPLRCOPGTRTOPRSHPAANDPSAAMSAAG ARGLRATYHRLLDKVELMLPEKLRPLYNHPA GPRTVFFWAPIMKWGLVCAGLADMARPAEK LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA VNFFVGAAGASQLFRIWRYNQELKAKAHK
929	2279		8007	2	1016	EFARRRVFIAAREMSLLRSLRVFLVARTGSYP AGSLLRQSPQPRHTFYAGPRLSASASSKELLM KLRKTGYSFVNCKKALBTCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ
930	2280	A	8008	3	1679	YVQPQGVSVVDFVRFECGEGEEAAETE  NSRVWGPWTEPSAGSI.RPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTTTLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP
931	2281	A	8009	861	300	AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL AEMGLAVDPNRAVPLRKRKVKAMEVDIEER

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			ŀ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
J				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ļ	İ			peptide	ĺ	/-possible nucleotide deletion, \-possible
<del></del>	<del></del>		<del> </del>	sequence	<del></del>	nucleotide insertion  PKELVRKPYVLNDLEAEASLPEKKGNTLSRD
						LIDYVRYMVENHGEDYKAMARDEKNYYOD
ľ		1		1	i .	TPKQIRSKINVYKRFYPAEWQDFLDSLQKRK
Į			ļ			MEVE TALL TRACTION OF THE WOOD LOSE OF T
932	2282	A	8011	412	1 i	SNLCLGNSWRWRWAKSRHHCIPTVTLSKRSG
ĺ		1				DIRGSHFSSPQRQRSQRVPGKETARVLRAGK
	}		1		ì	QGRGQIPIPCPWPPPPPPPPPGSPGPGCRQFHQ
			ł			SLEAKARHPASVREMRGKVKMRRALRRAPA
<u></u>						STRASSRQPNPK
933	2283	A	8012	147	1077	PPVPPASRSDMAQNLKDLAGRLPAGPRGMGT
	,	1				ALKLLLGAGAVAYGVRESVFTVEGGHRAIFF
		i	1 2		1	NRIGGVQQDTILAEGLHFRIPWFQYPIIYDIRA
		l				RPRKISSPTGSKDLQMVNISLRVLSRPNAQEL
			1			PSMYQRLGLDYEERVLPSIVNEVLKSVVAKF
						NASQLITQRAQVSLLIRRELTERAKDFSLILDD
		1	]			VAITELSFSREYTAAVEAKQVAQQEAQRAQF
		i		·		LVEKAKQEQRQKIVQAEGEAEAAKMLGEAL
	·					SKNPGYIKLRKIRAAQNISKTIATSQNRIYLTA DNLVLNLQDESFTRGSDSLIKGKK
934	2284	Α	8023	255	982	SQFSLSQVLVDSAEEGSLAAAAELAAQKREQ
						RLRKFRELHLMRNEARKLNHQEVVEEDKRL
i i		ł				KLPANWEAKKARLEWELKEEEKKKECAARG
		1				EDYEKVKLLEISAEDAERWERKKKRKNPDLG
						FSDYAAAQLRQYHRLTKQIKPDMETYERLRE
						KHGEEFFPTSNSLLHGTHVPSTEEIDRMVIDLE
			i l			KQIEKRDKYSRRRPYNDDADIDYINERNAKF
025	2007		0000	-50		NKKAERFYGKYTAEIKQNLERGTAV
935	2285	A	8027	59	310	LVSSTVNLLTEKAPWNSLAWTVTSYVFLKFL
						QGGGTGSTGMRDSALTLLGIGPSHRHSLSIRL
936	2286	A	8032	1	639	SQHSSPAPMYSQTFHILVLG
/30	2200	·	0032	•	039	SGRECNMAKTYDYLFKLLLIGDSGVGKTCVL FRFSEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ
			f (			IWDTAGQERFRTTITAYYRGAMGIMLVYDIT
1			.			NEKSFDNIRNWIRNIEEHASADVEKMILGNKC
						DVNDKRQVSKERGEKLALDYGIKFMETSAK
					1	ANINVENAFFTLARDIKAKMDKKLEGNSPOG
						SNQGVKITPDQQKRSSFFRCVLL
937	2287	Α	8039	393	311	EETIHSENSYILEKYIPISANLTLTIA
938	2288	A	8052	675	-1334	LHPAATSTAWLHVPPGLSMALSWVLTVLSLL
						PLLEAQIPLCANLVPVPITNATLDRITGKWFYI
				•		ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF
				1		LREYQTRQDQCIYNTTYLNVQRENGTISRYV
						GGQEHFAHLLILRDTKTYMLAFDVNDEKNW
	·			ļ		GLSVYADKPETTKEQLGEFYEALDCLRIPKSD
939	2289	A	8055	12	1039	VVYTDWKKDKCEPLEKQHEKERKQEEGES SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF
		4.	3033		1037	AEQLKWSAELARLGESIMDGKQGGMDGSKP
]	j		]		]	AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA
		İ				IHGHQLSLRNLISQGWAVNIITADHVSPLHEA
				ļ		CLGGHLSCVKILLKHGAQVNGVTADWHTPL
'						FNACVSGSWDCVNLLLQHGASVQPESDLASP
	-				1	IHEAARRGHVECVNSLIAYGGNIDHKISHLGT
	1			.		PLYLACENQQRACVKKLLESGADVNQGKGQ
						DSPLHAVARTASEELACLLMDFGADTOAKN
	1			l	1	AEGKRPVELVPPESPLAQLFLEREGPPSLMOL
	ļ			ļ	l	CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH
016						L
940	2290	A	8058	2	1203	KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI
						ARLRTMDLHLFDYSEPGNFSDISWPCNSSDCI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	1	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1	ļ	j	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l		peptide	· -	/=possible nucleotide deletion, \=possible
	<u> </u>		L	sequence		nucleotide insertion
	1					VVDTVMCPNMPNKSVLLYTLSFIYIFIFVIGMI
l		i i	į .	1		ANSVVVWVNIQAKTTGYDTHCYILNLAIADL
				1	<u> </u>	WVVLTIPVWVVSLVQHNQWPMGELTCKVTH
		1	•			LIFSINLFGSIFFLTCMSVDRYLSITYFTNTPSS RKKMVRRVVCILVWLLAFCVSLPDTYYLKT
l	'					VTSASNNETYCRSFYPEHSIKEWLIGMELVSV
	1					VLGFAVPFSIIAVFYFLLARAISASSDQEKHSS
		į				RKIIFSYVVVFLVCWLPYHVAVLLDIFSILHYI
	•				ĺ	PFTCRLEHALFTALHVTQCLSLVHCCVNPVL
						YSFINRNYRYELMKAFIFKYSAKTGLTKLIDA
				·		SRVSETEYSALEQSTK
941	2291	Α	8059	73	432	DMAGLMTIVTSLLFLGVCAHHIIPTGSVVLPS
	1	1				PCCMFFVSKRIPENRVVSYQLSSRSTCLKAGV
						IFTIKKGQQFCGDPKQEWVQRYMKNLDAKQ
942	2292	A	8067	278	1262	KKASPRARAVAVKGPVQRYPGNQTTC
772	2232	^	6007	2/0	1202	GGIGEIKQRPSCLGRCLDPSLSVLMNISLGLGS
l		į į				VFSAVISQKPSRDICQRGTSLTIQCQVDSQVT MMFWYRQQPGQSLTLIATANQGSEATYESGF
		]				VIDKFPISRPNLTFSTLTVSNMSPEDSSIYLCSA
İ						GRQGTYEQYFGPGTRLTVTEDLKNVFPPEVA
		i l				VFEPSEAEISHTQKATLVCLATGFYPDHVELS
		1 1				WWVNGKEVHSGVSTDPQPLKEQPALNDSRY
	í (	i i				CLSSRLRVSATFWQNPRNHFRCQVQFYGLSE
		f l				NDEWTQDRAKPVTQIVSAEAWGRADCGFTS
		ł l				ESYQQGVLSATILYEILLGKATLYAVLVSALV
943	2293	A	8070	1	879	LMAMVKRKDSRG
743	1 22,3		8070	•	0/7	MVKVVPATRGNLPRSQLTGTHQHCQPREPKI TASERLRRRPRATARLRAHAAPPEPPLAVFAP
						PSDRKELLALPVACDPVIASVMSWVQAASLI
						QGPGDKGDVFDEEADESLLAQREWQSNMQR
		ĺ				RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA
						EVILNYGRLRGTLSALLSWCHLHNNNSTLINK
!	1	1	: I			INNLLDAVGQCEEYVLKHLKSITPPSHVVDLL
	1	1	<u> </u>	1	ł	DSIEDMDLCHVVPAEKKIDEAKDERLCENNA
				i		EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK
944	2294	Α	8073	1	797	PKPHMDFGTDSQF ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK
		-	30.5	١ ١	131	MAATSGTDEPVSGELVSVAHALSLPAESYGN
1		1				DPDIEMAWAMRAMQHAEVYYKLISSVDPQF
			]	1		LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK
		. [	ļ			SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD
		, [	ĺ	ſ		CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA
		, 1	ŀ			VYISVQDKEGEKGVNNGGEKRADSGEEENT
		. !			1	KNGGEKGADSGEEKEEGINREDKTDKGGEK
945	2295		9074	<del></del>	506	GKEADKEINKSGEKAM
743	2673	A	8074	2	505	GAATLLRSASSAARKAAEAEQVWLHLHRYL
			j			SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT
-		. ]	.			GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW
		1	•	ĺ	ſ	GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER
		- :			1	ADLIAYLKKATNE
946	2296	A	8081	42	590	EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF
	ľ	į	1	ł	- 1	VAIFAVPLILGQEYEDEERLGEDEYYQVVYY
	1	)		i	1	YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK
		- 1	ļ		ļ	DITEALETTISLETARADHPKPVTVKPVTTEPQ
	. 1	- 1	}	J	J	SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC
947	2207	<del></del> _	9094	-200		KKVGRRLLMTLWMGVWQEEIGR
241	2297	A	8084	322	549	GGGSSPRELAGAAGLTVTSQAVAARRQQPSF
	· [				1	SRARAPAHSLRAALSLASSARSWGAVSRDRG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleo
948	2298	В	8093	3905	846	PCPPAIMYQSSNKC  MEPGEVKDRILENISLSVKKLQSYFAACEDEI PAIRNHDKVLQRLCEHLDHALLYGLQDLSSG YWVLVVHFTRREAIKQIEVLQHVATNLGRSR AWLYLALNENSLESYLRLFQENLGLLHKYYV KNALVCSHDHLTLFLTLVSGLEFIRFELDLDA PYLDLAPYMPDYYKPQYLLDFEDRLPSSVHG SDSLSLNSFNSVTSTNLEWDDSAIAPSSEDYD FGDVFPAVPSVPSTDWEDGDLTDTVSGPRST ASDLTSSKASTRSFTQRQNPFNEEPAETVSSS DTTPVHTTSQEKEEAQALDPPDACTELEVIRV TKKKKIGKKKKSRSDEASPLHPACSQKKCA KQGDGDSRNGSPSLGRDSPDTMLASPQEEGE GPSSTTESSERSEPGLLIPEMKDTSMERLGQPL SKVIDQLNGQLDPSTWCSRAEPPDQSFRTGSP GDAPERPPLCDFSEGLSAPMDFYRFTVESPST VTSGGGHHDPAGLGQPLHVPSSPEAAGQEEE GGGGGGGTPRPLEDTTREAQELEAQLSLVRE GPVSEPEFTQEVLCQLKRDQPSPCLSSAED GVDEGQGSPSEMVHSSEFRVDNNHLLLLMIH VFRENEBQLFKMIRMSTGHMEGNLQLLYVLL TDCYVYLLRKGATEKPYLVEEAVSYNELDY VSVGLDQQTVKLVCTNRRKQFLLDTADVAL AEFFLASLKSAMIKGCREPPYPSILTDATMEK LALAKFVAQESKCEASAVTVRFYGLVHWED PTDESLGPTPCHCSPPEGTTTKEGMLHYKAGT SYLGKEHWKTCFVVLSNGILYQYPDRTDVIP LLSVNMGGEQCGGCRRANTTDRPHAFQVILS DPPCLELSAESEAEMAEWMQHLCQAVSKGVI PQGVAPSPCIPCCLVLTDDRLFTCHEDCQTSF FRSLGTAKLGDISAVSTEPGKEYCVLEFSQDS QQLLPPWVIYLSCTSELDRLLSALNSGWKTTY QVDLPHTAIQEASNKKKFEDALSLIHSAWQR SDSLCRGRASRDPWC*
949	2299	A	8095	9	2374	ARRADTVLLESPSMLQGLLPVSLLLSVAVSAI KELPGVKKYEVVYPIRLHPLHKREAKEPEQQ EQFETELKYKMTINGKIAVLYLKKNKNLLAP GYTETYYNSTGKEITTSPQIMDDCYYQGHILN EKVSDASISTCRGLRGYFSQGDQRYFIEPLSPI HRDGQEHALFKYNPDEKNYDSTCGMDGVL WAHDLQQNIALPATKLVKLKDRKVQEHEKY IEYYLVLDNGEFKRYNENQDEIRKRVFEMAN YVNMLYKKLNTHVALVGMEIWTDKDKIKIT PNASFTLENFSKWRGSVLSRRKHDIAQLITA TELAGTTVGLAFMSTMCSPYSVGVVQDHSD NILRVAGTMAHEMGHNFGMFHDDYSCKCPS TICVMDKALSFYIPTDFSSCSRLSYDKFFEDKL SNCLFNAPLPTDIISTPICGNQLVEMGEDCDC GTSEECTNICCDAKTCKIKATFQCALGECCEK CQFKKAGMVCRPAKDECDLPEMCNGKSGNC PDDRFQVNGFPCHHGKGHCLMGTCPTLQEQ CTELWGPGTEVADKSCYNRNEGGSKYGYCR RVDDTLIPCKANDTMCGKLFCQGGSDNLPW KGRIVTFLTCKTFDPEDTSGEIGMVANGTKCG DNKVCINAECVDIEKAYKSTNCSSKCKGHAV CDHELQCQCEEGWIPPDCDDSSVVFHFSIVVG VLFPMAVIFVVVAMVIRHQSSREKQKKDQRP LSTTGTRPHKQKRKPQMVKAVQPQEMSQMK PHVYDLPVEGNEPPASFHKDTNALPPTVFKD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
050	0200		0100	peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
950	2300	A	8100	1	1251	MGLLMILASAVLGSFLTLLAQFFLLYRRQPE PPADEAARAGEGFRYIKPVPGLLLREYLYGG GRDEEPSGAAPEGGATPTAAPETPAPPTRETC YFLNATILFLFRELRDTALTRRWVTKKIKVEF EELLQTKTAGRLLEGLSLRDVFLGETVPFIKTI RLVRPVVPSATGEPDGPEGEALPAACPEELAF EAEVEYNGGFHLAIDVDLVFGKSAYLFVKLS RVVGRLRLVFTRVPFTHWFFSFVEDPLIDFEV RSQFEGRPMPQLTSIIVNQLKKIIKRKHTLPNY KIRFKPFFPYQTLQGFEEDEEHIHIQQWALTE GRLKVTLLECSRLLIFGSYDREANVHCTLELS SSVWEEKQRSSIKTGTISLTAVPMGWHRVSE AFPGLWYKLLVDLPFWGLEDGGPLLTVPLRQ CPG
951	2301	A	8108	1612	839	EVALFCFEMAAGMYLEHYLDSIENLPFELQR NFQLMRDLDQRTEDLKAEIDKLATEYMSSAR SLSSEEKLALLKQIQEAYGKCKEFGDDKVQL AMQTYEMYDKHIRRLDTDLARFEADLKEKQI ESSDYDSSSSKGKKKGRTQKEKKAARARSKG KNSDEEAPKTAQKKLKLVRTSPEYGMPSVTF GSVHPSDVLDMPVDPNBPTYCLCHQVSYGE MIGCDNPDCSIEWFHFACVGLTTKPRGKWFC PRCSQERKKK
952	2302	A	8112	595	291	PSVASLARRFSGRALWPPSHSVPGNRALCPRL LHGTTLPGGNQRELARQKNMKKQSDSVKGK RRDDGLSAAARKQRDSTPRDSEIMQQKQKK ANEKKEEPK
953	2303	A	8118		669	VCAGIRDPCSTPLAKPAAGGAENLSFGKQPG LETNILKMTTPNKTPPGADPKQLERTGTVREI GSQAVWSLSSCKPGFGVDQLRDDNLETYWQ SDGSQPHLVNIQFRRKTTVKTLCIYADYKSDE SYTPSKISVRVGNNFHNLQEIRQLELVEPSGW IHVPLTDNHKKPTRTFMIQIAVLANHQNGRD THMRQIKIYTPVEESSIGKFPRCTTIDFMMYRS IR
954	2304	A	8133		1015	PPLPPRSFPNLFSRPEPLPEPGRRGCNRSREPA ARAPSPPPPFEGAPGRAMVKVTFNSALAQKE AKKDEPKSGEEALIIPPDAVAVDCKDPDDVV PVGQRRAWCWCMCFGLAFMLAGVILGGAY LYKYFALQPDDVYYCGIKYIKDDVILNEPSAD APAALYQTIEENIKIFEEEEVEFISVPVPEFADS DPANIVHDFNKKLTAYLDLNLDKCYVIPLNT SIVMPPRNLLELLINIKAGTYLPQSYLIHEHMV ITDRIENIDHLGFFTYRLCHDKETYKLQRRETI KGIQKREASNCFAIRHFENKFAVETLICS
955	2305	A	8143	35	1171	VESRSAWHEGEDQIDRLDFTRNQMNILTTLDV KKKIKEVTEEVANKVSCAMTDEICRLSVLVD EFCSEFHPNPDVLKIYKSELNKHIEDGMGRNL ADRCTDEVNALVLQTQQEIIENLKPLLPAGIQ DKLHTLIPCKKFDLSYNLNYHKLCSDFQEDIV FRFSLGWSSLVHRFLGPRNAQRVLLGLSEPIF QLPRSLASTPTAPTTPATPDNASQEELMITTLVT GLASVTSRTSMGIIIVGGVIWKTIGWKLLSVS LTMYGALYLYERLSWTTHAKERAFKQQFVN YATEKLRMIVSSTSANCSHQVKQQIATTFARL CQQVDITQKQLEEEIARLPKEIDQLEKIQNNS KLLRNKAVQLENELENFTKQFLPSSNEES
956	2306	A	8157	1854	798	ASGSPAPSSSSAMAAACGPGAAGYCLLLGLH LFLLTAGPALGWNDPDRMLLRDVKALTLHY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, P=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  DRYTTSRILDPIPQLKCVGGTAGCDSYTPKVI QCQNKGWDGYDVQWECKTDLDIAYKFGKT VVSCEGYESSEDQYVLRGSCGLEYNLDYTEL GLQKLKESGKQHGFASFSDYYYKWSSADSC NMSGLITTVVLLGIAFVVYKLFLSDGQYSPPP YSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQ NTGHGATSGFGSAFTGQQGYENSGPGFWTGL
957	2307	A	8159	1492	528	GTGGILGYLFGSNRAATPFSDSWYYPSYPPSY PGTWNRAYSPLHGGSGSYSVCSNSDTKTRTA SGYGGTRRR THVVMTGMCYAPHQVLSYINGVTTSKPGVSL VYSMPSRNLSLRLEGLQEKDSGPYSCSVNVQ DKQGKSRGHSIKTLELNVLVPPAPPSCRLQGV PHVGANVTLSCQSPRSKPAVQYQWDRQLPSF QTFFAPALDVIRGSLSLTNLSSSMAGVYVCKA HNEVGTAQCNVTLEVSTGPGAAVVAGAVVG TLVGLGLLAGLVLLYHRRGKALEEPANDIKE DAIAPRTLPWPKSSDTISKNGTLSSVTSARAL RPPHGPPRPGALTPTPSLSSQALPSPRLPTTDG AHPQPISPIPGGVSSSGLSRMGAVPVMVPAQS
958	2308	A	8161	2340	1192	QAGSLV  ELARRPKQQSSEKSRNMIRNWLTIFILFPLKLV EKCESSVSLTVPPVVKLENGSSTNVSLTLRPP LNATLVITFETIFRSKNITILELPDEVVVPPGVT NSSFQVTSQNVGQLTVYLHGNHSNQTGPRIR FLVIRSSAISIINQVIGWIYFVAWSISFYPQVIM NWRRKSVIGLSFDFVALNLTGFVAYSVFNIGL LWVPYIKEQFLLKYPNGVNPVNSNDVFFSLH AVVLTLIIIVQCCLYERGGQRVSWPAIGFLVL AWLFAFVTMIVAAVGVTTWLQFLFCFSYIKL AVTLVKYFPQAYMNFYYKSTEGWSIGNVLL DFTGGSFSLLQMFLQSYNNDQWTLIFGDPTK FGLGVFSIVFDVVFFIQHFCLYRKRPGYDQLN
959	2309	A	8163	521	1345	GERAGRRGRLGVWAQPQPLLPRPVGSRRE MQPPGPPPAYAPTNGDFTFVSSADAEDLSGSI ASPDVKLNLGGDFTKESTATTFLRQRGYGWL LEVEDDDPEDNKPLLEELDIDLKDTYYKIRCV LMPMPSLGFNRQVVRDNPDFWGPLAVVLFFS MISLYGQFRVVSWITTWIFGSLTTFLLARVLG GEVAYGQVLGVIGYSLLPLIVIAPVLLVVGSF EVVSTLIKLFGVFWAAYSAASLLVGEEFKTK KPLLTYPIFLLYTYFLSLYTGV
960	2310	A	8167		2921	MTCFKGQKGEQRSHAFEANKDHKAKVPSPN LYSQLNALQFTVDERSILWLNQFLLDLKQSL NQFMAVYKLNDNSKSDEHVDVRVDGLMLK FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSQFWIDYEGM KSGKGRPISFVDSPPLSIWICQPTRYAESQKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI HLLVHVHKHVSMQINHYQYLLLLFLHESLILL SENLRKDVEAVTGSPASQTSICIGILLRSAELA LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN GDFLSSKRKQISRDINRIRSVTVNHMSDNRSM SVDLSHIPLKDPLLFKSASDTNLQKGISFMDY LSDKHLGKISEDESSGLVYKSGSGEIGSETSD KKDSFYTDSSSVLNYREDSNILSFDSDGNQNI LSSTLTSKGNETIESIFKAEDLLPEAASLSENL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN			F-rienymannic, G-Grycine, H-Histidine,
seq-	uence	]		location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	ucitot	i i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ł		Į		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	i	i		peptide		/=possible nucleotide deletion, \=possible
Ĺ				sequence		nucleotide insertion
					<u> </u>	DISKEETPPVRTLKSQSSLSGKPKERCPPNLAP
1		İ				LCVSYKNMKRSSSQMSLDTISLDSMILEEOLL
I	ļ	1	l	1		ESDGSDSHMFLEKGNKKNSTTNYRGTAESVN
1		Ι.		ĺ	}	AGANLQNYGETSPDAISTNSEGAQENHDDLM
				1	1	SVVVFKITGVNGEIDIRGEDTEICLQVNQVTP
	l		l			DQLGNISLRHYLCNRPVGSDQKAVIHSKSSPE
1	1	1			ł	ISI DEEGCOCA VIIIGI I A EKNIGEI OGUTTO TOTA
i	ĺ		ļ			ISLRFESGPGAVIHSLLAEKNGFLQCHIENFST
l .	l		į į			EFLTSSLMNIQHFLEDETVATVMPMKIQVSNT
	i .			1		KINLKDDSPRSSTVSLEPAPVTVHIDHLVVER
<b>!</b> .	l				i	SDDGSFHIRDSHMLNTGNDLKENVKSDSVLL
					}	TSGKYDLKKQRSVTQATQTSPGVPWPSQSAN
			]			FPEFSFDFTREQLMEENESLKQELAKAKMAL
100	0011					AEAHLEKDALLHHIKKMTVE
961	2311	Α	8172	1442	682	TAAMSIFTPTNQIRLTNVAVVRMKRAGKRFEI
			[			ACYKNKVVGWRSGVEKDLDEVLQTHSVFVN
Ì						VSKGQVAKKEDLISAFGTDDQTEICKQILTKG
			<u> </u>	1		EVQVSDKERHTQLEQMFRDIATIVADKCVNP
1						ETKRPYTVILIERAMKDIHYSVKTNKSTKQQA
						LEVIKQLKEKMKIERAHMRLRFILPVNEGKKL
1			l			KEKLKPLIKVIESEDYGQQLEIVCLIDPGCFREI
			1			DELIKKETKGKGSLEVLNLKDVEEGDEKFE
962	2312	A	8175	286	587	NISNKAEVSSHPSVISHSMDSFGQPRPEDNQS
		••	0.75		307	VLRRMQKKYWKTKQVFIKATGKKEDEHLVA
						CDAELDANI EVERICIOETOTELI NITURNO D
}						SDAELDAKLEVFHSVQETCTELLKIEKYQLR
963	2313		0101	12	0016	LNGMKS
763	2313	Ÿ.	8181	13	2215	AEGCAERRGTEPVVELSMSWESGAGPGLGSQ
						GMDLVWSAWYGKCVKGKGSLPLSAHGIVV
1	i					AWLSRAEWDQVTVYLFCDDHKLQRYALNRI
		•				TVWRSRSGNELPLAVASTADLIRCKLLDVTG
						GLGTDELRLLYGMALVRFVNLISERKTKFAK
1 1						
1 1						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI
l i					.	VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI
					. •	
						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE
						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK
						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL
						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA
-						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL
						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE
-						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT
•						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP
•						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD
						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK
						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ
•						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE
						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVOPFSTGQESPTA
						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEBNVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF
						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEBNVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV
						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEBNVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF
	•					VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQUEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVOFFSTGGESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF
964	2314	A	8184	6	1393	VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQUEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVOFFSTGGESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF
964	2314	A	8184	6	1393	VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEFEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRTRGRRRRACRSAE
964	2314	A	8184	6	1393	VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRTRGRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR
964	2314	A	8184	6	1393	VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRTRGRRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP
964	2314	Α	8184	6	1393	VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERAREILVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIFKAMQQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTILGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRTRGRRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFFFCTIDPNESR
964	2314	A	8184	6	1393	VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQCEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTILGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFFFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG
964	2314	A	8184	6	1393	VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA
964	2314	A	8184	6	1393	VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRTRGRRRACRSAE GTGLRSLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI
964	2314	A	8184	6	1393	VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERAREILVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPTTRGRTRGRRRACRSAE GTGLRSLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS
964	2314	A	8184	6	1393	VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERAREILVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFSEAKAQQEEQ GSVNDVKEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVOPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPTTRGRTRGRRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFFFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS WVIDQKKPVRFYHDWNDKEIEVLNKHLFLTS
964	2314	A	8184	6	1393	VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERAREILVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPTTRGRTRGRRRACRSAE GTGLRSLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS
964	2314	A	8184	6	1393	VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERAREILVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFSEAKAQQEEQ GSVNDVKEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVOPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPTTRGRTRGRRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFFFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS WVIDQKKPVRFYHDWNDKEIEVLNKHLFLTS

NO: of   No: of   bod	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
peptide coticle sed uses of correspondity to the correspondity of the correspondity of peptide sequence prepride sequence of peptide seq			hod		beginning	•	D=Aspartic Acid. E=Glutamic Acid.
Decision   Decision	nucl-	peptide	İ	in			
Sequence   1944   1944   1945   194	eotide	seq-	}	USSN	location	corresponding	
967 2317 A 8210 3 601 Sawdessesses and residue of peptide residue of peptide sequence pepti	seq-	uence		09/496	correspondi		M=Methionine, N=Asparagine, P=Proline
### squence   sq	uence			914			O=Glutamine R=Arginine S=Serine
residue of peptide sequence   Y=Tyrosnia, X=Uaktaown, =-Stop codon,   Y=Stop c	1		·				T=Threonine V=Valine W=Tryptophan
peptide /-possible nuclotide deletion, \( \to \) possible nuclotide insertion   nuclotid	[ .			İ	residue of		Y=Tyrosine X=IInknown *=Ston codon
Sequence	}		ļ	}		]	Enossible nucleatide deletion Impossible
ANMIGSALPKIKKGFAALQLEYFFTAGPIE RAWTIKGTKAPQAGKHITIPKEGFIMAE MKYEDFKERGSENA VKAAKGYRQQGRNYIT EDGDIJEFKINTPOOPPKK  SPESTILSFELLSFESSMALGAAGATRVFYAM AALGGPILGVSATLINSVILSSAIKSIN PPP GGAAGIFICSAVSAAPGIL YPGGNKYQTIDN QPYPCAEDECQTIBCYCASPTROGRACYOG LACRKERKCMEHAMCCPGNYCKINGICVSS DONINFRGEIEETTESFONDHISTLOVISRETT LSSKMYHTIKGGRSVCLRSSDCASGLCAER FWSKLKVILKEGOVCTKIFRKGSKGSKGLEIFO RCYCGIGI SCRIQKDHIQASNISKLIHTOPRE PWSKLKVILKEGOVCTKIFRKGSKGSKLIHTOPRE RCYCGIGI SCRIQKDHIQASNISKLIHTOPRE RCYCGIGI SCRIQKDHIQASNISKLIHTOPRE RCYCGIGI SCRIQKDHIQASNISKLIHTOPRE RCYCGIGI SCRIQKDHIQASNISKLIHTOPRE RCYCGIGI SCRIQKDHIQASNISKLIHTORI RCYCGIGI SCRIQKDHIQASISKLIHTORI RCYCGIGI SCRIQKDHIQASNISKLIHTORI RCYCGIGI SCRIQKDHIQASISKLIHTORI RCYCGIGI SCRIQKDHIQASISCHIMI RCYCGIGI SCRIQKDHIQAGASISKLIHTORI RCYCGIGI SCRIQKDHIQAGASISKLIHTORI RCYCGIGI SCRIQKDHIQ	1		[				
RAWTIEKGTKAPQAAGKEHTDFEKGFIMAEV MKYEDFKEGGENAVKAAGKYQQGRNYT EDGDIFFKFNIFQOPKKK  965 2315 A 8195 1437 594 RSPSISSILSPSEMMALGAAGATRVFVAM AAALGGIPLLGVSATLNSVIAIKNIJSNAIKNLPPP GGAAGHEGSAVSAAPGILYPGGKKYQTIDN QPYPCAEDEEGGTDEYCASPTRGGDAGVQIG LACRKRRKCMRHAMCCPGNYCKGIGCVS DQNIFFRGEHEITTESFONDHSTLDGVSRRTI LSKMYHTKGQRGSVCLRSDCASGLCCAR FWSKICKPVLKEGQVCTHERKGSHGLEIP RCYCGGGLSCEIQKNUHQASNSSKHITOQRR RCYCGGGLSCEIQKNUHQASNSSKHITOQRR RCYCGGGLSCEIQKNUHQASNSSKHITOQRR RCYCGGGLSCEIQKNUHQASNSSKHITOQRR RCYCGGGLSCEIQKNUHQASNSSKHITOQRR RCYCGGGLSCEIQKNUHQASNSSKHITOQRR RCYCGGGLSCEIQKNUHQASNSKHITOQRR RCYCGGGLSCEIQKNUHQASNSKHITOQRR RCYCGGGLSCEIQKNUHQASNSKHITOQRR RCYCGGGLSCEIQKNUHQASNSKHITOQRR RCYCGGGLSCEIQKNUHQASNSKHITOQRR RCYCGGGLSCEIQKNUHQASNSKHITOQRR RCYCGGGLSCEIQKNUHQASNSKHITOQRR RCYCGGGLSCEIQKNUHQASNSKHITOQRR RCYCGGGLSCEIQKNUHQASNSKHITOQRR RCYCGGGLSCEIQKNUHQASNSKHITOQRR RANTINGORYSTERITAGRASHTOQRR RANTINGORGCGITVTDMKGNNSH LLSALKYPANVA'DEPVERFINSSENGSICGGCQCQP SPONSWHITOQRRASHTINGORGCGITVTDMKGNNSH LLSALKYPANVA'DEPVERFINSSEVAGSLY, RADLDOVGKALLETSERTIAVSLOULDKRI FWQYNRBGSNSLICSCDIYDGGSVHISHIPTC RCACHARASTGGDRFYSTWKMKTIVANNCHTC KDMYRRHASSFOPICTSTWKMKTIVANNCHTC KDMYRRHASSFOPICTSTWKMKTIVANNCHTC KDMYRRHASSFOPICTSTWKMKTIVANNCHTC KDMYRRHASSFOPICTSTWKMKTIVANNCHTC KDMYRRHASSFOPICTSTWKMKTIVANNCHTC KDMYRRHASSFOPICTTCSTWCTCTC RCGCCGCSSPONGGGGGLCCYP SPVSWECDCPFGYDLQLEBSCAASGOPQFT LANNGQDRAMFTGSYCTCTCFFULPT GRACHQLYBCALVEWARANDGSQ RELLEEGGVYCCCSGCSPONGGGGLCKRIPACCC RCGCRMKASDURKTTCKDRVLQSSMKRSLVVIPI LANGGQVACCSGNYARCSSGGDATCOCLA RCGCRMASSGCGGHICKKRICATAWCS CREGFMKASDGKTCLADDGQLGVIPISDA RRLTCRDDVCHPFAVACCCVCTCMCARLSSEGDATCOCLA RCGCNASCTTRGGGYTCMCAGRLSSEGDATCOCLA RCGCNASCTTRGGGYTCMCAGRLSSEGDATCOCLA RCGCNASCTTRGGGYTCMCAGRLSSEGDATCOCLA RCGCNASCTTRGGGYTCMCAGRLSSEGDATCOCLA RCGCNASCTTRGGGYTCMCAGRLSSEGDATCOCLA RCGCNASCTTRGGGYTCMCAGRLSSEGDATCOCLA RCGCNASCTTRGGGYTCMCAGRLSSEGDATCOCLA RCGCNASCTTRGGGYTCMCAGRLSSEGDATCOCLA RCGCNASCTTRGGGCTHCXSSKCOTT GGGYYCCCSGGYGGGGHICKKRICATA RCGCYPACCSGYGGGHICKKRICATAWCS CREGFMKASDGCTCAACCC RCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC					Doquence	<del> </del>	
MKYEDFKERGSENAVKAAGKYRQQGRNTI EDGDIPFKINTPOPPKK  8578LSYBLISSEMMALGAAGATRVYVAM AALGGPILGVSKATINSVINSNIANINIPPP GGAAGHPGSAVSAAPGILYPGGNKYQTIDN QPYPCAEDEGGTDEVCASPTRGGDAGVQIG LACRKRRKCMRHAMCCPGNYCKNIGICVS DQNIFFGBEEGTTESFGNBHSTLOGYSRTT LISSKMYHTKGQBGSVCLRSSDCASGLCCAR FWSKLCKPVLKEGQVCTKHRKKGSHGLEIQ RCYCGEGLSCKIQKDHHQASNSSRLHTQQRR FWSKLCKPVLKEGQVCTKHRKKGSHGLEIQ RCYCGEGLSCKIQKDHHQASNSSRLHTQQRR RCYCGEGLSCKIQKDHHQASNSSRLHTQQRR RCYCGEGLSCKIQKDHHQASNSSRLHTQRR WYDLERQLLQRVFLNGSRQERVCHIEKNVSS AANINMINEEVIWENQGGGITYTMKGRNSH LLSALKYPANVAVDPVERFIFWSSEVAGSLY RADLDGVGWKALLSTESKTAVSLVDLKKR FWQYNREGSNSLCSCDYDDGGSVHISKHIPTQ RADLDGVGWKALLSTESKTAVSLVDLKKR FWQYNREGSNSLCSCDYDDGGSVHISKHIPTQ RADLDGVGWKALLSTESKTAVSLVDLKKR FWQYNREGSNSLCSCDYDDGGSVGOLQSHL MCAEGYALSRDKXCEGODWKYCEDVINCE AFWHGCTLGCKNTPGSYVCTCPGVFULLPI GKRCHOLVSCPRNVSECSHDCVLTSEGPLCP CPEGSVLERDGKTCSGCSSPDNGGSQLCVP SPYSWEDCPGPGVJLQLBEKSCAASGPQPFL LANSQDRMMHPGTDYGTLSQQNGMVY ALDHDPVENKTYFAHTALKWERANDGSQ RELIEEGVDYPEGLAVDWGRRPYWTDRGK SLIGRSDLAGKGSKTVGGCSGLCVR SPYSWEDCPGVJJQLGBKAANDGSQ RELIEEGVDYPEGLAVDWGRRPYWTDRGK SLIGRSDLAGKGKTUCHASDLAW RRFUTTOGDFRESSSLQGGRUVFWSDWAAN SVIRVNKTTCKDRVRLQCSSMLKPSSLVVUB LAKFGADCLYCHGOGGGHCCDBCCQGVBS CREGGWAASDCKTCLALDGHQLLAGGGYDL KWQTTLDILSTRYGCGCGHICKKRLGTA WCS CREGGWAASDCKTCLALDGHQLLAGGGYDL KWQTTLDILSTRYGCDCHCCDBCCQLGVBS CGRAASTTHTGGYTCMCAGRLSERGLICTD BGGYVCRCSSGYOGGGCHCCDBCCQLGVBS CGRAASTTHTGGYTCMCAGRLSERGLICTD LKWWBLRTGAGGQCAVIVAVCVVLV MLLLISLWGAHYYKTCKLLSKNFKNFYEESS RDVRSRRADTBEDGMSSCPQPWFVVCPASSKCGT BGGYVAGEDQAADGSMQFTSWGCEGQCVVVAVCVVLV MLLLISLWGAHYYKTCKLLSKNFKNFYEESS RDVRSRRADTBEDGMSSCPQPWFVVCPASSKCGT BGTGGGCVVFSSKGSCCQVVVECHOD LKWGGGPVAGEDQAADGSMQFTSWGCEGQ LCGMGTGGGCVTVSSKGSCCQVVECHEDD LKWGGGPVAGEDQAADGSMQFTSWGCEGG LCGMGTTTTTGGGYCCTRSSLKGCCQCQCVVVALVECHOD LKWGGGPVAGEDQAADGSGCGGGAACAGGASLCCCAGASCCAACGCCAACCACACACACACACACACACACA	ł l					1	D A DATE TO A DO A A CHIEF DEEP CED A DU
P65   2315   A   8195   1437   594   SESSISSISSISSISSISSISSIMALIGAGATRIVEVAM   AAALGGIPLLGVSATILNSVLNISNALKIL.PPP   GAAAGHPCSAVAAPGIL.YPGEKYYTIDM   QPYPCABDEECGTDEYCASPTRGGDAGVIC   LACKRERKERMEIAMCCPONYCKNIGLVS   DONIFIRGEIEETTTESPONDISTILDGYSEKRYTITL   LSEKMYLITKGGEOSVCLESBOCASGLCCARI   FWSKICKPVILKEGOVCTKERKKGSHOLEIPP   RCYCGEGLSKORICHHAASNSKILHTCORE   FWSKICKPVILKEGOVCTKERKKGSHOLEIPP   RCYCGEGLSKORICHHAASNSKILHTCORE   RCYCGEGLSKORICHHAASNSKILHTCORE   RCYCGEGLSKORICHHAASNSKILHTCORE   RCYCGEGLSKORICHHAASNSKILHTCORE   RCYCGEGLSKORICHHAASNSKILHTCORE   RCYCGEGLSKORICHHAASNSKILHTCORE   RCYCGEGLSKORICHHAASNSKILHTCORE   RCYCGEGLSKORICHHAASNSKILHTCORE   RCYCGEGLSKORICHHAASNSKILHTCORE   RCYCGEGLSKORICHHAASNSKILHTCORE   RCYCGEGLSKORICHHAASNSKILHTCORE   RCYCGEGLSKORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNSKORICHTCORE   RCYCGEGLSKORICHHAASNORICHHAASSILHE   RCYCGEGLSKORICHHAASSILHE   RCYCGEGLSKORICHTCORE   RCYCGEGLSKORICHHAASSILHE   RCYCGEGLSKORICHTCORE   RCYCGEGLSKORICHHAASSILHE   RCYCGEGLSKORICHTCORE   RCYC	1 1		ł	1		1	
Post   Post		,					
AAALGGEPLLGVSATLNSVLNSNAUKLEPPP GGAAGHPGSAVSAAPGILYPGGNKYQTIDH QPYPCAEDEEGOTDEYCASTRIGGDAGVQIG LACKKRKKRCMRIAMCCPGNYCKNGIGCY DANIFRGEEETITESFONDHSTLDGYSRKTT LSSKMYHTKGGBGSVCLESSDCASGLCAG FWSKICKPVLEGGOVCTKHRKGSHGLEIPQ RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF PWSKICKPVLEGGOVCTKHRKGSHGLEIPQ RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF DTEGTNYEGLVVDAGVSVIMDFHYNEKRIY WVDLERGLAGRVFLNGSRQERVCNEKNIVS MAINWINEEVIWSNQQEGIITVIMGNNSH LLSALKYPANVAVDPVERFIFWSSEVAGSLY RADLDGVGVKALLETSEKTAVALLVILKON RDTWEPEQKLCKLRKGNCSSTVCGQDLQSHISKHPTC HNLFAMSLRGDRIFYSTWKKKTIWIANKHTIC KDMYRRILSISSPVLGELXVYHACKTOWLANKHTIC KDMYRRILSISSPVLGELXVYHACKTOWLANKHTIC KDMYRRILSISSPVLGELXVYHACKTOWLANKHTIC KDMYRRILSISSPVLGELXVYHACKTOWLANKHTIC KDMYRRILSISSPVLGELXVYHACKTOWLANKHTIC KDMYRRILSISSPVLGELXVYHACKTOWLANKHTIC KDMYRRILSISSPVLGELXVYHACKTOWLANKHTIC KDMYRRILSISSPVLGELXVYHACKTOWLANKHTIC GKRCHQLVSCFRNYSECSHDCVLTSECPFLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCASGCVPY SPVSWEDCDFTGYDLQLDEKSGASGCVPY LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHIDPVBNKYTGYLQLAGKSAGPOPFL LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHIDPVBNKYTGYLQLAGKAGAGPOPFL LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHIDPVBNKYTGYLAGKAGAGCCR REGENWASCSGTOWRYSCSHDAGAGAGCC REGENWASCSGTOWRYGCAGAGCCC REGENWASCSGTOWRYGCAGAGCCC REGENWASCTGRORYTGAGAGCCC REGENWASCTGRORYTGAGAGCCC REGENWASCTGRORYTCAGAGCCC REGENWASCTGRORYTCAGAGCCCC REGENWASCTGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	965	2315	Δ	2105	1427	504	DCDCI CECL COOPE O (A) CA CACO COOPE
GGAAGHPGSAVSAAPGILYPGGNKYQTION  PYPYCABDEECTOFBYCASTRGDAGVQIC  LACKERKRCMRIAMCCPGNYCKNGICY  DANHFRGEIESTTESPONDHSTLOYSRETT  LSSKMYHTKGGBGSVCLRSSDCASGLCACH  FWSKICKPYLKEGQVCTCHRRKGSHGLEIPQ  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RCYCGEGLSCRIQKDHHQASNSSRLHTCQRIF  RADLDGYGVKALLETSEKTIAVSLDVILKGRINS  RADLDGYGVKALLETSEKTIAVSLDVILKGRINGS  RADLDGYGVKALLETSEKTIAVSLDVILKGRINGS  RADLDGYGVKALLETSEKTIAVSLDVILKGRINGS  RADLDGYGVKALLETSEKTIAVSLDVILKGRIP  HNLFAMSLGREYSTYMKMKTIWIANKHTI  KDMYRRINLISSFYLGGLCXVILTSEGPILCF  GKRCHQLVSCFRINSECSIDLCVILTSEGPILCF  GKRCHQLVSCFRINSECSIDLCVILTSEGPILCF  GKRCHQLVSCFRINSECSIDLCVILTSEGPILCF  GKRCHQLVSCFRINSECSIDLCVILTSEGPILCF  GKRCHQLVSCFRINSECSIDLCVILTSEGPILCF  GKRCHQLVSCFRINSECSIDLCVILTSEGPILCF  GKRCHQLVSCFRINSECSIDLCVILTSEGPILCF  GKRCHQLVSCFRINSECSIDLCVILTSEGPILCF  GKRCHQLVSCFRINSECSIDLCVILTSEGPILCF  GKRCHQLVSCFRINSECSIDLCVILTSEGPILCF  GKRCHQLVSCFRINSECSIDLCVILTSEGPILCF  GKRCHQLVSCFRINSECSIDLCVILTSEGPILCF  GKRCHQLVSCFRINSECSIDLCVILTSEGPILCF  RCYCRCSGGYGGLGHLADGROCHARCKART  REFITION BY THE TOTAL THE TOTA	1 ,03	2.713	Λ	6173	1457	J <del>34</del>	
966 2316 A 8207 416 4082 KFKLEKMELANCOSYSTEMELSPONDEN CONTROLOGICAN PROSECUTION OF THE PRINCIC PVILKE GOVERNOON CONTROLOGICAN PROSECUTION OF THE PRINCIC PVILKE GOVERNOON OF THE PRINCIP PVINCIC PVILKE GOVERNOON OF THE PRINCIP PVINC	1 1					ĺ	
LACRKRRKCMRHAMCCPONYCKNIGICYS   DONHTRIGEIETTIESSPONTSTLDGYSRRTI   LSSKMYHTKGQPGSSYCLRSSDCASGLCAR   FWSKICKPVLKEGQVCTRKSGHOLEIPO    RCYCGEGLSCRIQKDHHQASNSSRLHTCQRF   RCYCGEGLSCRIQKDHHQASNSSRLHTCQRF   RCYCGEGLSCRIQKDHHQASNSSRLHTCQRF   RCYCGEGLSCRIQKDHHQASNSSRLHTCQRF   RCYCGEGLSCRIQKDHHQASNSSRLHTCQRF   RCYCGEGLSCRIQKDHHQASNSSRLHTCQRF   RCYCGEGLSCRIQKDHHQASNSSRLHTCQRF   RCYCGEGLSCRIQKDHHQASNSSRLHTCQRF   RCYCGEGLSCRIQKDHHQASNSSRLHTCQRF   RCYCLRCRICTION   RCYCLRCRICTION     WYDLRRQLLQRVFLINGSQRCVCNIEKNVSC   MAINWINEEVINSNQQEGUITVTDMKGNNSH   LLSALKYPANAVDPVPERIFWSSEVAGSLY,   RADLDGVGYKALLETSEKTAVSLUVLDKRL   FWIQYNREGSISLICSCDYDGGSVHISKIPTIC   HOLFAMSL FGDRIFYSTWKMKTTWIANKHTC   KDMVRNIH INSSYPLICEL LYDINGSPYLICELONEC   APWRIGCILGCKNTFGSYYTCTCVGFVLLIFI   GKRCHQLVSCTRNVSSCSHDCVLTSEPICF   CPEGSVLERBGRTCSGCSSPVGGCSQLCVFI   GKRCHQLVSCTRNVSSCSHDCVLTSEPICF   CPEGSVLERBGRTCSGCSSPVGGCSQLCVFI   GKRCHQLVSCTRNVSSCSHDCVLTSEPICF   CPEGSVLERBGRTCSGCSSPVGGCSQLCSCROPPIC   LFANSQDERMHFDGTTOTICLSQQMGMYY   ALDHDPVENKIYFAHTALKWIERANMOSQ   RERLIEBGGPUYFEGLAV DWIGKTSSTLLSQQMGMYY   ALDHDPVENKIYFAHTALKWIERANMOSQ   RERLIEBGGPUYFEGLAV DWIGKTSSTLLSQQMGMYY   ALDHDPVENKIYFAHTALKWIERANMOSQ   RERLIEBGGPUYFEGLAV DWIGKTSSTLLSQQMGMYY   ALDHDPVENKIYFAHTALKWIERANMOSQ   RERLIEBGGPUYFEGLAV DWIGKTSYTLTSGQMGMCYTAMSSDLMW    RCHTTTANAS LARGE BEATCACLK   RCHTTTANAS LARGE BEATCACLK   KNQVTPLDILSKTRVSEDNITESQHMLVAEIM   VSDQDDCAPYCCSMYARCISGEBDATCQCLK   KNQVTPLDILSKTRVSEDNITESQHMLVAEIM   VSDQDDCAPYCCSMYARCISGEBDATCQCLK   GFAGBGKLCSDIDECEMGYPVCPPASSKCINT   EGGYVCRCSBGYQGDGHCKSGGEDATCQCLK   GFAGBGKLCSDIDECEMGYPVCPPASSKCINT   EGGYVCRCSBGYQGDGHCKSGGEDATCQCLK   STPPPHLREDDHHYSVRNSDEORYPTENSELWM   STRANGRSSEGGAPTTMEGQGVTCMCAGRIFFICHCD   STPPPHLREDDHHYSVRNSDEORYPTHSQHLVAEIN   GEGYVCRCSBGYQGDGHCKSGGEDATCQCLK   GRASSCTNTEGGGRTCKCHCHCR   GRASSCTNTEGGGRTCKCHCHCR   GRASSCTNTEGGGRTCKCHCHCR   GRASSCTNTEGGGRTCKCHCHCR   GRASSCTNTEGGGRTCKCHCHCR   GRASSCTNTEGGGRTCKCHCHCHCHCHCHCHCHCHCHCHCHCHCHCHCHCHCHC							GGAAGHPGSAVSAAPGILYPGGNKYQTIDNY
P66 2316 A 8207 416 4082 KFKLIKRMCGGSVCLKSSDCASGLCARS FYSKICKPVLKEGGVCTKHRRKGSHGLEIPQ RCYCGEGLSCRQKDHHQASNSRLHTCQRE RCYCGEGLSCRQKDHHQASNSRLHTCQRE SCRUCKDHLQASNSRLHTCQRE SCRUCKDHLQASNSRLHTCQRE SCRUCKDHLQASNSRLHTCQRE SCRUCKDHLQASNSRLHTCQRE WYDLERQLLQRVFLNGSRQGRVCNIERVISK MAINWINEDLYVDAGVSVIMDFHYNEKRIY WYDLERQLLQRVFLNGSRQGRVCNIERVISK MAINWINEDLYVDAGVSVIMDFHYNEKRIY WYDLERQLLQRVFLNGSRQGRVCNIERVISK MAINWINEDLYVDAGVSVIMDFHYNEKRIY WYDLERQLLQRVFLNGSRQGRVCNIERVISK MAINWINEDLYVDAGVSVIMDFHYNEKRIY WYDLERQLLQRVFLNGSRQGRVCNIERVISK AMAINWINEDLYVDAGVSALLETTSERLYSLDVLDAKR LLSALK VPANVA DPVPERIFWSSEVAGSL Y RADLDGVGVALLETTSERLYSL HILFAMSLFGDRIPYSTWKKKTWIANKHTC KDMVRIN-HSSFYPLGELKVVHPLAQPKAED DTWEFEQKLCKLRKGNCSSTVCGQDLQSHLM MCAGGYALSRDRKVCEGDVNEC AFWNIGCTLGCKNTPGSYYCTCPVGFVLLPI GKRCHQLVSCPRNVSECSHOWNYCEDVNEC AFWNIGCTLGCKNTPGSYYCTCPVGFVLLPI GKRCHQLVSCPRNVSECSHOWNYCEDVNEC CPEGSVLERDGRTCSGCSSPDNGGCSQLCVP SPVSWECDCFPGVDLQTHSEPDLCF GKRCHQLVSCPRNVSECSHOWNYCEDVNEC AFWNIGCTLGCKNTPGSYYCTCPVGFVLLPI GKRCHQLVSCRASGPOPPL LFANSQDERMHPDGTDYGTLLSQQMGMYV ALDHDPYEDKRYTPATALKWIERANMCDSQ REPLIEEGVDVPPGCLAVDWIGRRPYWTDRGSC SLIGRSDLNGKRSSITIENISQPRGJAVPMAR RLFWTDTGINFRESSLQGRILVASSDLLW PSGTTDFLTDKLTWCDAKQSVLEMANLDGSS RRRLTQNDVGHPPAVAVPEDYWFSDWAMD SVIRVNKRTGKDRVRLQGSMLKPSSLVVVIP LAKRGADPCLYONGCEBLKKRIGTJA WCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDLSKRTQTACGCCRASGTVGDCHGTCAGGCTYR LAKRGADPCLSCPGVT GGGAGACTCAGGRACHT CGCC CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDLSKRTVSEDIDTESGCPCYR GGGAGACTCAGRACHT GGGAGACTCAGACCC CREGFMKASDGKTCLALDGHQLAGGEVDL KNGQVYACEDQAADGSCPQVWFFVXVKEHQD LKNGGVYVACEDQAADGSCPQVWFFVXWEHQD LKNGGVYVACEDQAADGSGCPQVWFFVXWEHQD LKNGGQVYACEDQAADGSCPQVWFFVXWEHQD LKNGGQVYACEDQAADGSCPQVWFFVXWEHQD LKNGGQVYACEDQAADGSCPQVWFFXWQEGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC							QPYPCAEDEECGTDEYCASPTRGGDAGVQIC
LISSKMYHTKGQGGSVCLKIRSRGGSIGCARI   FWSKICKPVLKEGQVCTKIRRKGSIGCARI   FWSKICKPVLKEGQVCTKIRRKGSIGCARI   RCYCGGGLSCRIQKDHQASNSSRLHTCQRI   RCYCGGGLSCRIQKDHQASNSSRLHTCQRI   RCYCGGGLSCRIQKDHQASNSSRLHTCQRI   RKKLRMLTLILLPVVSKFSFYSLSAPQHW   SCPEGTLAGNGNSTCVGAPFLIFSHGMSIFRI   DTEGTNYEQLVVDAGVSVIMDHYYMSKRIY   WVDLERQLLQNFULNGGGSVTUDHKYMSKRIY   WVDLERQLLQNFULNGGGSVTUDHKYMSKRIY   WVDLERQLLQNFULNGKNSH   LLSALKYPANAVDPVERFIFWSSEVAGSLY   RADLDGVGVKALLETSEKTAVSLDVLDKRI   FWIQYNREGSISLICSCDYDHAQPKARD   DTWEPEQKLCKLRKGNCSSTVCGQDLQSHL   MCAEGYALSROPKYCEGDVNEC   AFWNHGCTLGCKNTPGSYYCTCPVGFVLLPY   GKRCHQLVSCPRNVSECSHCCVLTSEGPLCP   CPEGSVLERRGKTSGCSSPNGGCSQLCVPY   SPYSWECDCFPGYDLQLDEKSCAASGQPPL   LFANSQDIEMHHDGTDYGTLAQGMGMYV   ALDHDPVENKTYAHTALKWIERANMOSQ   RERLIEGGVDVPEGLAVDWIGRRYWTDRGK   SLIGRSDLNGKRSKITTENISQPRGIAVIPMAR   RLFYDTJGINPRIESSSLQGLIVASSDLW    PSGTIDFLTDKLYWCDAKOSVIEMANLDGSR   RRRLTQRDVGHPRAVAFEDYVWFSDWAMI   SVIRVNKRTGKDRVRLQGSMLKPSSLVVHP   LAKFGADPCLYQNGGGEHICKKRLGTAWCS   CREGFMAASDGKTCALLODHQLLAGGEVDL   KNQVTPLDLLSKTRVSEDNITESQHRUVAEM   VSDQDDCAPYCCSMYARCISGGEDATCQCLK   GFAGDGKLCSDIDECEMGVPVCPPASSKCINT   EGGTVCRCSSGYQGGGHICKRLGTAWCS   CGENASCTNTEGGTVCACAGRISGEGLCD   STPPPHLRIDDHTYSVRNSDGCPSH-SHOPYCL   GEGYVCRCSSGYQGGGHICKSRLGTLAGGE   DLKWWELRHAGHGQQKVTVAVCVVVLV   MLLLSLWGAHYYKTQKLISAPPLWPYESS   RDVSRRPADITEDGMSSCPQWFVVKEHQD   LKNGGQVACEDQAADGSCPQFFVVWEHQD   LKNGGQVACEDQAADGSCPQFFVWKEHQD   LKNGGQVACEDQAAGSCPQFFVWKEHQD   LKNGGQVACEDQAAGSCPQFFVWKEHQD   LKNGGQVACEDQAAGSCPQFFVWKEHQD   LKNGGQVACEDQAAGSCPQFFVWKEHQD   LKNGGQVACEDQAAGSCPQFFVWKEHQD   LKNGGQVACEDQAAGSCPQFFVWKEHQD   LKNGGQVACEDQAAGSCPQFFVWKEHQD   LKNGGQVACEDQAAGSCPQFFVWCHCHQD   LKNGGQVACEDQAAGSCPQFFVWCHCHQD   LKNGGQVACEDQAAGSCPQFFVWCHCHQD   LKNGGQVACEDQAAGSCPQFFVWCHCHQD   LKNGGQVACEDQAAGSCPQFFVWCHCHQD   LKNGGQVACEDQAAGSCPQFFVWCHCHQCAACD   DPHOMELTQ   LCHARTSTALAFRY	!!						
FWSKICKPVLKEGQVCTKHRKKGSHGLEIPQ RCYCGEGLSCRIQKDHOASNSSLHTCQRE RCYCGEGLSCRIQKDHOASNSSLHTCQRE SCPEGTLAGNGNSTCVGPAPFLIFSHGNSIFRI DTEGTNYEQLVVDAGVSYIMDPHYNEKRY WVDLERQLLQRVPLNGSRGERVCNIEKNYSG MAINWINERUWSNQGGIITVTDMKGNNSH LLSALKYPANVAVDPVERFIFWSSEVAGSLY RADLDGVGVALLETSEWKMKTTWIANKHTC KDMYRINLHSSPVLGELKVVHPLAQPKAED DTWEPEQKLCKKGNCSSTVCGDLQSHLK MCABGYALSRDRKYCEGDNWKYCEDVHSCH AFWNHGCTLGCKNTPGSYYCTCYGFVLLPH GRKCHQLVSCPRNVSECSHDCVLTSEGPLCF CPEGSYLERDGKTCSGCSSPDNGGSQLCVPH LFANSQDIRHMHFDGTDYGTLLSQMGMVY ALDHDPVENKIYFAHTALKWIERANMDGSQ RERLEEGVDVPBGLAVDWIGREPYWTDRGK SLIGRSDLNGKRSKITTENISQPRGIAVHYMAG RRLTQNDTGHPFAVAVFEDYVWFSDWAM SVIRVNKRTGKDRVRLQGSMLKPSLVVHP LAKPGADPCLVYONGGCHICKKRIGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDIJSKTRVSEGGIHCLDIDECQLGVHS GGYVCRCSEGGYGGGHICLDIDECQLGVHS GGYVCRCSEGGYGGIHCLDIDECQLGVHS GGYVCRCSEGGYGGIHCLDIDECQLGVHS GGYVCRCSEGGYGGGHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHTSVRNSDSECPLSHDGYCL HIGGVCMYTEALDKYACNCVVVILV MLLILSLWGAHTYRTOKLISKNPKRDFYEESS RDVRSRRPADTEDDMSCCPQWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQFTSWRQEFQ LCGMGTEQGCWPVSSDKGSCPQVMERSFH MPSYGTGTLEGGVERPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGRSSSHAAVIPDGDSIRRETGFSQASLL RLHHRPRALDRNKKGYLSRRMDLQQGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKFEHNSRRNKLHYAPQLY DLDRDGKISRHEMIQVLRLMVGQVTEEQL ENADDTVQAADEGDGQAVSFVEFTKSLEKM	1 1						DQNHFRGEIEETTTESFGNDHSTLDGYSRRTT
RCYCGEGI.SCRIQKDHHQASNSSR.HT/CQRE    SCREEN   REALKMILTHILLIPUVSKESFYSI.SAPQHW    SCREGTLAGNCRISTCVGPAPPLIFSHGNISTR    DTEGTNYEQLVVDAGVSVIMDFHYNEKRIY    WYDLERQLLQRVFLINGSRQERVCNIERNYS    MAINWINEEVIWSNQQEGIITYTDMGNINSH    LLSALKYPANVAVDPVERFIFWSSEVAGSLY    RADLDGVGVKALLETSEKTIAVSLUVLDKR, I FWIQYNIRGSINSLICSCOPPOGSSVHISKHPTC    HILFAMSLFGDREYSTWKMKTTWANKHTC    KDMVRINLHSSFVPLGELKVVHPLAQPKAED    DTWEPEQKLCKLRKGNCSSTVCGDDLQSHLA    MCAEGYALSRDKKYCEGNDWKYCEDVNEC    AFWNHGGTLGCKNTPGSYCTCPVGFVLLPI    GRRCHQLVSCPRNVSECSHDCVLTSEGPLCF    CPEGSVLERDGKTCSGCSSPDNGGCSQLCVP    SPVSWECDCFFGYDLQLDESSCAASGPQPFL    LPANSQDIRHMIFDGTDYGTLLSQQMGMVY    ALDHDPVENKIYFAHTALKWIERANMOGSQ    RERLIEBGVDVPBGLAVDWIGRFYWTDRGK    SLIGRSDLNGKRSKITTEINSQPRGIAVHPMAE    RLFWTDTGRIPRIESSSLQGLGRLIASSDLLW    PSGTTIDFLTDKLYWCDAKGSVIEMANLDGSS    RRRLTQNDVGHPFAVAVFEDYWFSDWAM    SVRVNKRTGKDRVRLQGSMLKPSSLVVHP    LAKPGADPCLYQNGGCEHICKKRLGTAWCS    CREGFMKASDGRTCALDGHQLLAGGEVDL    KNQVTPLDIJSKTRVSEDNITESQHMLVAEM    VSDQDDCAFVGCSMYARCISEGBDATCQCLK    GFAGDKLCSDIDECEMVPVCYPASSKCDNT    EGGYVCRCSEGVGGDGHCLDIDECOLGVHS    CGEMASCTNTTEGGYTCMCAGRLSEPGLLCPD    STFFPHLREDDHHYSYRNESECPLSEIDGYCL    HIGHOVCMYIEALDKYACNCVGVIGERCQYR    DLKWWELRHAGHGQQGVIVVAVCVVVIV    MLLLISLWGAHYYRTOKLLISKNPNYYESS    RDVFSRRPADITEDGMSSCPQPWFVVIKEHQD    LKNGGQPVAGEDQQAADGSMQPTSWRQEPQ    LCGMGTEQGCWIPVSSDKGSCPQWFVVIKEHQD    LKNGGQPVAGEDQQAADGSMQPTSWRQEPQ    LCGMCTEQGCWIPVSSDKGSCPQWFVVIKEHQD    LKNGGQPVAGEDQQAADGSMQPTSWRQEPQ    LCGMCTEQGCWIPVSSDKGSCPQWFVVIKEHQD    LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ    LCGMCTEQGCWIPVSSDKGSCPQWFVVIKEHQD    LKNGGQPVAGEDGQAADGSWGPTSWRQEPQ    LCGMCTEQGCWIPVSSDKGSCPQWFVVIKEHQD    LKNGGQPVAGEDGQAADGSWGPTSWRQEPQ    LCGMCTEQGCWIPVSSDKGSCPQWFVVIKEHQD    LKNGGQPVAGEDGQAADGSWGPTSWRQEPQ    LCGMCTEQGCWIPVSSDKGSCPQWFVVILCHQLC    ENADDTVQBAADGAAOSFVFITKSLEKM    VEDEDTETQDKKRETHLSRRNKLHYAPQLY    DLDRBGSRRHEMIQVURLANGGVQVTEEQL    ENADDTVQBAADGAAOSFVFITKSLEKM	l i	-					LSSKMYHTKGQEGSVCLRSSDCASGLCCARH
See							FWSKICKPVLKEGQVCTKHRRKGSHGLEIFQ
SCPEGTLAGNGNSTCVGPAPFLIPSHGNSIFRI DTEGTNYEQLVVDAGVSVIMDHYNEKRIY WYDLERQLQRVFLNGSRQERVCNIEKNYS MANWINEEVIWSNQGEGITVTIMKGNINSH LLSALKYPANAVDPPERFIFWSSEVAGSLY RADLDGVGVKALLETSEKITAVSLDVLDKR, FWIQYNREGSISHLISCDFGSVHISKIPTO HNLFAMSLFGDRIPYSTWKMKITWIANKHTG KDMVRINLHSSFVFLGELKVVHFLAGKAED DTWEPEQKLCKLRKGNSTVCGODLOSHL MCABGYALSRDRKYCEGNDWKYCEDVINEC AFWNHGGTLGCKHTPGSYVCTCPVGFVLLPI GRKCHQLVSCPRNVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPI SPVSWECDCFPGYDLQLDEKSCAASGPOPFL LFANSQDIRHMHFDGTDYGTLLSQMGAVYY ALDHDPVERNKIYAAHTALKWIERANMDGSQ RERLIEBGVDVPSGLAVDWIGRFAYWTDRGK SLIGRSDLNGKSKITIEINSQPRGLAVIPMAB RLFWTDTGINFRIESSSLQGLGRLVASSDLIW PSGTTDFLTDKLYWCDAKQSVIEMANLDGSS RRAITONDVGHPFAVAVFEDYVWFSDWAMI SVIRVNIKRTGKDRVRLQGSMLKRIGTAWCS CREGFMKASDGRTCLALDGHQLLAGGEVDL KNQVTFLDILSKTRVSEDHITESQHMLVAEIM VSDQDDCAFVGCSMYARCISEGBDATCQCLK GFAGDGKLCSDIDECEMPVCPPASSKCENT BGGYVCRCSEGVQGDGHCLDIDECOLGVHS CGENASCTNITEGGYTCMCAGRLSEPGLICPD STPPHILEBDHHYSVRNSDSECPLSHDGYCL HIGGVCMYIBALDKYACNCVVGVIVGERCQYT DLKWWELRHAGHGQQCKIVVAVCVVVIV MILLISLWGAHYYRTIGKLLSKNPKNPYEESS RDVPSRRFADITEOMSSCPOWFWVIKEHOD LKNGGQPVAGEDQQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSINGSCPOYWENSHE MSYGTQTLEGGVGRVFSSLSAANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQGALLAV NPILGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDFKKPETJNSRRKKLHYAPQLY DLDRDGKISRHEMIQVLRLMVGQVTEEQQL ENMADRTVQAADEGDGAAVSFVEFTKSLEKM	066	2216		0207	416	4000	KCYCGEGLSCRIQKDHHQASNSSRLHTCQRH
DITEGTINYEQLLYVDAGYSVIMDPHYNEKRIY WYDLERQILQRYFLINGSRQER VCNIERCHYSS MAINWINEEVIWSNQQEGIITVTDMKGNNSH LLSALKYPANVAVDPVERIFIPKSSEVAGSLY RADLDGYGVKALLETSEKITAVSLDVLDKRL FWIQYNREGSISLICSCDYDGGSVHISKHPITT HINLFARSL FGDRIFYSTWKMKTIWIANKHTI KDMYRINLHSSFYPLGELKYVHPLAQPKAED DTWEPEQKLCKLRKGNICSSTVCQGDLQSHL MCAEGYALSRDRKYCEGRIDWKYCEDVNEC AFWNHIGCTLGCKNTPGSYYCTCPVGFVLLPI GRCHQLVSCFRNYSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPI SPYSWECDCFPGYUL DIRESCAASGPQPRL LFANSQDIRHMHPDGTDYGTLLSQQMGMVY ALDHDPVENNITYAHTALKWIERANMDGSQ REALEEGVDVPEGLAVDWIGRRFYWTDRCK SLIGRSDLNGKRSKITTEINSQPRGJAVHYMAR RLFWTDTGINFRIESSLQGLGRLVTASSDLIW PSGITDFLTDKLJWCDAKQSVEMANLDGSK RRILTQNDVGHPFAVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVRLQGSMLKYSSLVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLAUGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMY ARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHICLIDECQLGYMI STPPHLREDDHHYSVRNOESCPLSHDGYCL HDGVCMYMFALDKYACNCVGYGIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLISLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADITEDGMSSCPQWFVVKEHQD LKNGGQFVAGEDGQAADGSMQTTSWRQEFQ LCGMGTGQGCWPVSSDKGSCPQVWFVKEHQD LKNGGPVAGEDGQAADGSMQTTSWRQEFQ LCGMGTGQGCWPVSSDKGSCPQVWFVKEHQD LKNGGPVAGEDGGAADGSMQTTSWRQEFQ LCGMGTGQGCWPVSSDKGSCPQVWFVKEHQD LKNGGPVAGEDGGAADGSMQTSWRQEFG SAMMSRSSHAAVIPDGDSIRRETGFSQASIL RLHHRRALDRNKKGYLSRMDLQQIGALAV NPLGDREUBSFPPDGSQDFGVFVLHATHEP VEDEDTETQDPKKPEPLNSRRNKLHYAPQLY DLDRDCKISRHEMLQVLRILMVQQVTEELD LENLADRTVQEADEDGDGAVSPVEFTKSLEKM	200	2010	A	6207	416	4082	KFKLIKIMLLTLIILLPVVSKFSFVSLSAPQHW
WYDLERQIL QRYFLNGSRQER VCNIERNYSS MANWINEEVIWNQCGGUITYTDMKGNNSH LLSALKYPANVAVDPVERFFWSSEVAGSLY RADLDGYGVKALLETSEKITAVSLDVLDKRIL FWQYNREGSNSILGSCDYDGGSVHISKHPTC HNLFAMSLFGDRIFYSTWKMKTIWIANKHTC KDMYRINLHSSFYPLGENVYHLAQPKAED DTWEPEQKLCKLRKGNCSSTVCGQDLQSHLA MCABGYALSRDRKYCEGNDWKYCEDVHEC AFWHINGCTLGCKNTTGSTYCTCPVGFVLLPI GKRCHQLVSCPRNVSECSHDCVLTSEGPLCF CPGSVLREDGKTCSSPDNIGGCSQLCVPI SPVSWECDCFPGYDLQLDEKSCAASGPQPFL LFANSQDIRHMHFDGTDYGTLJSQQMGMVYY ALDHDPVENKTYFAHTALKWIERANMOSSQ RERLIEEGYDVPEGLAVDWIGGRFY WTDRGK SLIGRSDLINGKRSKITTENISOPRGIA-VIPMAR RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSS RRRLICTONDVCHPFAVAPEDYVWFESDWAM SVIRVNKRTGKDRVRLQGSMLKPSSLVVVIP LAKPGAPCLYONGCHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRV SEDNITESQHMLVAEIM VSDQDDCAPYGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCDTI EGGYVCRCSEGYQGGICLDDDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPHLREDDHHYSVRNDSSECLSIDGYGL HDGVCWYMFALDRYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVIV MLLLSLUGAHYYRTOKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LIKNGGQPVAGEDQGAADGSMQFTSWRQFFG UCGMGTGQGCWIPVSSDKGSCPQVWRQFSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTO  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRRALDRNKKGYLSRMDLQQIGALAV NFLGDRIEBSFPPOLOGDFFGVYRUTAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAPQLY DLDRDKISREPMLQVILRIMVQQVYTEQL ENHADRTVQEADEDGDGAVSFVEFTKSLEKM	Î						
MAINWINEEVIWSNQQEGIITYTDMKGNNSH LISALKYPANVAVDPVERFIFWSSEVAGSLY RADLDGVGVKALLETSEKITAVSLDVLDKRI FWQYNREGSINSLICSCDYDGGSVHISKHPIT HNLFAMSLFGDRIPYSTWKMKITWIANKHTG KDMVRINLHSSFVPLGELKVVHPLAQPKAED DTWEPEQKLCKLRKGNCSSTVCGQDLQSHL MCABGYALSRDRKYCEGRDWKYCEDVNEC AFWNHGCTLGCKNTPGSYYCTCPVGFVLPI GRRCHQLVSCPRNVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDMGCSQLCVPI SPVSWEDCDFPGYDLQLDERSCAASGPQPL LFANSQDIRHMHPDGTDYGTLLSQQMGMVYY ALDHDPVENKTYPAHTALKWIERANMDGSQ RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLNGKRSKITTENISGPRGIAVHPMAR RLFWTDTGINFRIESSLQGEGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSS RRRLICONOVGHPEAVAVEDYVWFSDWAMI SVIRVNKRTGKDRVRLQGSMLKPSSLVVHP LAKPGADPCLYQNGGCBHICKRRIGTAWGS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESGHMLVAEIM VSDQDDCAPVGCSMYARCISGEBDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHCLIDDEQLGVFD STPPPHLREDDHHYSVRNDGSECPLSHDGYCL HDGVCMYMEALDKYACNCVVGVIGERCQYR DLKWWELRHAGHQQQKVIVVAVCVVVLV MLLLSLWGAHYYRTQKLLSKNRKNPYESG ROWGWFAGDGGAADGSMQFTSWRQEFQ LCGMGTGQGCWPVSDKGSCCPQWFVVIKEHQD LKNGGPVAGEDGGAADGSMQFTSWRQEFQ LCGMGTGQGCWPVSDKGSCCPQWFVVIKEHQD LKNGGPVAGEDGGAADGSMQFTSWRQEFG LCGMGTGQGCWPVSDKGSCCPQWFVVIKEHQD SAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRRALDRNKKGYLSRMDLQQIGALAV NPLGDRIUGESFFPDGSGDFPGFVAVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAPQLY DLCRDGKISRHEMLQVLRILMVGQVTLEGU ENMADRTVQEADEDGJGAVSFVEFTKSLEKM	1						
LLSALKYPANVAVDPUERIFWSSEVAGSLY RADLOGGVYKALLETSEKITAVSLDVLDKRL FWIQYNREGSISLICSCDYDGGSVHISKHPIC HINLFAMSLFGDRIFYSTWMKTIWIANKHTG KDMVRINLHSSFVPLGELKVVHPLAQPKAED DTWEPEQKLCKLRKGRCSSTVCGQDLQSHL MCABGYALSRDRRYCEGNDWKYCEDVNEC AFWNHGCTLGCKNTPGSYYCTCPVGFVLLPI GKRCHQLVSCPRIVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPI SPVSWECDCFFGYDLQLDEKSCAASGPPEL LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHDPVENKITFAHTALKWIERANNDGSQ REBLIEEGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLINGKRSKITTENISQPRGIAVHPMAR RLFWTDTGRIPRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLLOGS RRRLTQNDVGHPFAVAVFEDYVWFSDWAMI SVRVNKRTGKDRVRLQSSMLKFSSLVVHP LAKPGADPCLVONGGCSMLKFSSLVVHP LAKPGADPCLVONGGCSMLKFSSLVVHP LKPGADPCLVONGGCSMLKFSSLVVHP SGFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCCSGGYGGGHCLKDREOCLGVHS GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCCSGGYGGGHCLDIBCCQLGVHS CGENASCTNTBGGYTCMCAGRLSEPGLICPD STPPHILREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYEBALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKTVVAVCVVVLV MLLLIS.WGAHPYRTQKLLSKNPKDYPEGS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQFTSWRQEPQ LCMMTEGGCBWPVSDKGSCPQVWERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ LCMGTEGGCWPVSDKGSCPQVWERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ LCMGTEGGCWPVLAHFREP VEDEDTETQDPKKPPLNSRRNKLHYAPQLY DLDRDGKISRHEMI_QVLRILMVGVQVTEEQL ENIADRTVQEADEDGGAVSVVEFTKSLEKM	1 1			1			
RADLDGVGVKALLETSEKITAVSLDVLDKRI FWIQYNREGSNSILSCCDYDGGSVHISKHPTO HNLFAMSLFGDRIFYSTWKMKITWIANKHTO KDMVRIN.HSSPYPLGELKVVHPLAQPKAED DTWEPEQKLCKLEKGNCSSTVCGDLQSHLL MCAEGYALSRDRKYCEGDVSHCCGDLQSHLL MCAEGYALSRDRKYCEGDWKYCEDVNEC AFWNHGCLGCKNTPGSYYCTCPVGFVLLPI GKRCHQLVSCPRIVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPI SPVSWECDCFPGJQLDERSCAASGPOPFI LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHDPVENKIYFAHTALKWERANNDGSQ RERLIEEGVDVPBCLAVDWIGRERYWTDRGK SLIGRSDLNGKRSKITTENISQPRGLAVHPMAE RLFWTDTGINFRIESSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVLEMANLDGSS RRRLTQNDVGHFFAVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVQGSMLKPSSL.VVVIPP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVYPLDILSKTRVSEDNITESGHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAAGGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHHCLDIDECQLGVHS CGENASCTNTBGGYTCMCAGRLSEPGILCPD STPPPHLREDDHIFYSVRNSDSECFLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKINPYEESS RDVRSRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDQQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQYMERSHH MPSYGTQTLEGGVERPYSSDKGSCPQYMERSHH MPSYGTQTLEGGVERPYSSDKGSCPQVMERSHH MPSYGTQTLEGGVERPYSSDKGSCPQVMERSHH MPSYGTQTLEGGVERPHSLSANPLWQQRAL DPPHQMELTQ LCGMGTEQGCWIPVSSDKGSCPQVMERSHH MPSYGTQTLEGGVERPSSRNKLHYAFQLY DLDRDGRIESFFPDGSQRVDFFGFWVLAHFIPP VEDEDTETQDPKKPPENSRRNKLHYAFQLY DLDRDGRIESFFPDGSQRVDFFGFWVLAHFIPP VEDEDTETQDPKKPPENSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMWGVQVTEEQL ENIADRTVQEADEDGGAVSFVFFTKSLEKM	1 1						
FWIQYNREGSNSLICSCDYDGGSVHISKHPTC HNLFAMSLFGDRIFYSTWKMKTTWIANKHTG KDMVRINLHSSFYPLGELKVVHPLAQPKAED DTWEPEQKLCKLRKGNCSSTVCGQDLQSHL MCABGYALSRDRKYCEGMDWKYCEDVNEC AFWNHGCTLGCKNTPGSYYCTCPVGFVLLPT GKRCHQLVSCPRIVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSCPINGGCSQLCVPI SPVSWECDCFPGYDLQLDEKSCAASGPQPFL LFANSQDIRHMHFDGTDVTLLSQQMGMVY ALDHDPVENKITYAHTALKWIERANNDGSQ RERLIEEGVDVPEGLAVDWIGREFY WTDRGK SLIGRSDLNGKRSKITTIENISQPRGIAVHPMAE RLFWTDTGNPRESSSLQGLGRLVIASSDLIW PSGITTDFLTDKLYWCDAKQSVIEMANLDGSR RRRLTQNDVGHPSESSSLQGLGRLVIASSDLIW PSGITTDFLTDKLYWCDAKQSVIEMANLDGSR RRRLTQNDVGHPSSSSLQGLGRLVAGSDLIV PSGITTDFLTDKLYWCDAKQSVIEMANLDGSR RRRLTQNDVGHPSSSSLVVVHP LAKFGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISGEDATCQCLK GFAGDGKLCSDIDECEMGYPVCTPASSKCITH EGGYVCRCSEGYQGDGHHCKBRLDDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHITYSVRNSDSECPLSHDGYQCL HDGVCMYIEALDKYACNCVVYQYIGERCQYR DLKWWELRHAGDQQKVTVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQFWFVVKEHQD LKNGGQPVAGEDQQAKOSMOPTSWQEFG LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ LCGMGTEGGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLSANPLWQQRAL DPPHQMELTQ LCGMGTEGGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPPLNSRRNKLHYAPQLY VEDEDTETOJPKKRPPLNSRRNKLHYAPQLY VEDEDTETOJPKKRPPLNSRRNKLHYAPQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENADDRTVQEADEDGGAVSFVEFTKSLEKM	{						
HNLFAMSLFGDRFYSTWKMKTIWIANKHTU KDMYRINLHSSPULGELKVVIPILAQPKAED DTWEPEQKLCKLRKGNCSSTVCGQDLQSHLG MCAEGYALSRDRKYCEGNDWKYCEDVNEC AFWNHGCTLGKNTFGSYYCTCPVGFVLLPI GKRCHQLVSCPRNVSECSHDCVLTSEGPLCF CPEGSVLERDGKTSCSGCSSPDNGGCSQLCVPI SPYSWECDCFPGYDLQLDEKSCAASGPQFFL LFANSQDRHMHFDGTDYGTLLSQQMGMVYV ALDHDPVENKTYSHTALKWERANNDGSQ RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLNGKRSKITTIENISQPRGIAVHPMAR RLFWTDTGINFSSSLQGGRLVVASSDLIW PSGTTDFLTDKLYWCDAKQSVIEMANLDGSR RRRLTQNDVGHPFAVAVFEDYVWFSDWAMI SVIRVNKRTGKFRVAVFEDYVWFSDWAMI SVIRVNKRTGKFRVAVFEDYVWFSDWAMI SVIRVNKRTGKFRVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQIMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEFGLICPD STPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYEALDKYACNCVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MILLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVSSRRPADTEDGMSSCPQFWFVIKEHQD LKNGGQPVAGEDQQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLSANPLWQQRAL DPPHQMELTQ  967 2317 Å 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRINKKGYLSRMDLQQIGALAV NPLGDRIESFIPGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAPQLY DLDRDGKISRHEMLQVLRLMVGVVVTEQL ENNADRTVQEADEDGGAVSFVEFTKSLEKM							
KDMVRINLHSSFYPLGELKVVHPLAQPKAED DTWEPEGKLCKLRKGNCSSTYCGODLQSHLG MCABGYALSRDRKYCEGNDWKYCEDVNEC AFWNHGCTLGCKNTPGSYYCTCPVGFVLLEP GKRCHQLVSCRNYSCSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPI SPVSWEDCPFGYDLQLDEKSCAASGPQFPL LFANSQDIRHMHPDGTDYGTLLSQQMGMVY ALDHDPVENKTYPAHTALKWERANMDGSQ RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLNGKRSKITTEINSQPRGIAVHPMAE RLFWTDTGINPRIESSSLQGLGRLVVASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPRAVAVFEDYVWFSDWAMM SVIRVNKRTGKDRVRLQGSMLKPSSLVVVIP LAKPGAPDCLYONGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDLSKTRVSEGHDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCNYIEALDKYACNCVGYIGERQQYR DLKWWELRHAGHGQQQKUTVAVCVVVLV MLLLISLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQFWFVVKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEFQ LCGMGTEQGCWFPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 Å 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAPQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGAOSSFVEFTKSLEKM							FWIQYNREGSNSLICSCDYDGGSVHISKHPTQ
DTWEPEQKLCKLRKGNCSSTVGGQDLQSHLA MCABGYALSRDRKYCEDNUMCYCEDVNEC AFWNHGCTLGCKNTPGSYYCTCPVGFVILPI GKRCHQLVSCPRNVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPI CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPI SPVSWECDCFPGYDLQLDEKSCAASGPQPFL LFANSQDRHMHPGTDYGTLLSQQMGMVY ALDHDPVENKIYPAHTALKWIERANMDGSQ RERLIEBGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLNGKRSKITTENISQPRGIAVHPMAR RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSR RRRLTQNDVGHPPAVAVEDYYWPSDWAMM SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDLSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISSGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPFASSKCINT EGGYVCRCSEQYOGDGHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STFPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYEBALVYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MILLLSLWGAHYYKTDYKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSKGSCPQVMERSFH MPSYGTQTLEGGVEPTKSLLEKM NPLGDRIESFFPDGSQRVDFPGFVRVLAHTRP VEDEDTETQDPKKFGYLLSMVGVQVTEEQL ENIADRTVQEADEDGGDGAVSFVEFTKSLEKM	[ ]						HNLFAMSLFGDRIFYSTWKMKTIWIANKHTG
MCAEGYALSRDRKYCEGNDWKYCEDVILER AFWNHGCTLGCKNTTPGSYYCTCPVGFVLLPI GKRCHQLVGSCPRINVSECSHDCVLTSEGFLCF GKRCHQLVGSCPRINVSECSHDCVLTSEGFLCF CPEGSVLERDGKTCSGCSSPDINGGCSQLCVPI SPVSWEGDCFGYDYDLQLDEKSCAASGPQPFL LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHDPVENKLYFAHTALKWIERANMDGSQ RERLIEEGYDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLNGKRSKIITIENISQPRGIAVHPMAR RLFWTDTGINPRIESSLQGLGRLVIASSDLIW PSGTIDFILTDKLYWCDAKQSVIEMANLOGSR RRRLTQNDVGHPFAVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVBLQGSMLKPSSLVVVHF LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISGEBDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQDGDIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSSCPQSHDGYGCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MILLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSTRRETGFSQASIL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPILGDRIESSFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY UDLDDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM	[						KDMVRINLHSSFVPLGELKVVHPLAQPKAED
AFWHIGCTLGCKNTPGSYYCTCPVGFVLLTB GKRCHQLVSCPRNVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSSPDNGGCSQLCVPI SPVSWECDCFFGYDLQLDER.SCAASGPQFPL LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHDPVENKIYFAHTALKWIERANMDGSQ RERLIEEGVDVPEGLAVDWIGRRPYWTDRGK SLIGRSDLNGKRSKITTENISQPRGIAVHPWTDRGK SLIGRSDLNGKRSKITTENISQPRGIAVHPMAR RLFWTDTGINFRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTFLDILISKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCYPASSKCINT EGGYVCRCSEGYQGDGHCLDDECQLGVHS GCENASCTNTEGGYTCMCAGRLSEPGLICPD STPPHLREDDHHYSVRNSDSESPLSHDGYQL HDGVCMYEALDKYACNCVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVLV MILLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVPSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDQQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFYRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM	ł ł						
GKRCHQLVSCPRNVSECSHDVJLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPI SPVSWEDCFPGYDLQLDEKSCAASGPQPFL LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHDPVENKIYFAHTALKWIERANMDGSQ RERLIEEGVDVPEGLAVDWIGREFYWTDRGKS SLIGRSDLNGKRSKIITIENISQPRGIAVHPMAR RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAM SVIRVNKRTGKDRVRLQGSMLKPSSLVVVIPP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHCLDIDECQLGVHS CGENASCTINTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRPRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTTETQDPKKPEPLNSRRNKLHYAPQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM	]						
CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPI SPVSWECDCFPGYDLQLDEKSCAASGPOFFL LFANSQDRHMHFDGTDYGTLLSQQMGMVY ALDHDPVENKIYFAHTALKWIERANMDGSQ RERLIEEGVDVPEGLAVDWIGRFFYWTDRGK SLIGRSDLNGKRSKITTENSQPPGIGAVHPMAR RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITIDFLIDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMN SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAFVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCTPASSKCINT EGGYVCRCSEGYQDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HIDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDQQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSDKOSCPQVMERSH MPSYGTQTLEGGYEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASIL RLHHRPRALDRIKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTIETQDPKKPEPLNSRRNKLHYAPQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM							AFWNHGCTLGCKNTPGSYYCTCPVGFVLLPD
SPYSWECDCFFGYDL QLDEKSCAASGPQFFL LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHDPVENKIYFAHTALKWIERANMDGSQ RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLINGKRSKITTISINSQPRGIAVHFMAR RLFWTDTGINPRIESSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHFPAVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILISKTRVSEDNITESQHMLVAEIM VSQDDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHICLDIDECQLGVHS CGETYCRCSEGYQGDGHICLDIDECQLGVHS CGETYCRCSEGYQGDGHICLDIDECQLGVHS CGETYCRCSEGYQGDGHICDDECQLGVHS CGETYCRCSEGYQGDGHICDDECQLGVHS CGETYCRCSEGYQGDGHICDDECQLGVHS CGETYCRCSEGYQGDGHICDDECQLGVHS CGETYCRCSEGYQGDGHICDDECQLGVHS CGETYCRCSEGYQGDGHICDDECQLGVHS CGETYCRCSEGYQGDGHICDDECQLGVHS CGETYCRCSEGYQGDGHICDDECQLGVHS CGETYCRCSEGYQGDGHICDDDECQLGVHS CGETYCRCSEGYQGDGHICDDDECQLGVHS CGETYCRCSEGYQGDGHICDDDECQLGVHS CGETYCRCSEGYQGDGHICDDDECQLGVHS CGETYCRCSEGYQGDGHICDDDECQLGVHS CGETYCRCSEGYQGDGHICDDDECQLGVHS CGETYCRCSEGYQGDGHICDDDECQLGVHS CGETYCRCSEGYQGDGHICDDDECQLGVHS CGETYCRCSEGYQGDGHICDDDECQLGVHS CGETYCRCSEGYQGDGHICDDDCQLGVHS CGETYCRCSEGYQGDGHICDDDCQLGVHS CGETYCRCSEGYQGDGHICDDCGCUGVHS CGETYCRCSEGYQGDGHICDDCGCGCUGVHS CGETYCRCSEGYQGDGHICDDCGCGCUGTSHI CHARACTOR CARACTOR CREGTYCRCSCATACAC CREGTYACACAC CREGTYCRCSCATACAC CREGTYACACAC CREGTYACACAC CREGTYACACACAC CREGTYCRCSCATACACAC CREGTYACACACACACACAC CREGTYACACACACACACACACACACACACACACACACACACAC	i i						
LFANSQDIRHMIFDGTDYGTLLSQQMGMVY ALDHDPVENKIYFAHTALKWIERANMDGSQ RERLIEEGVDVPEGLAVDWIGRFPYWTDRGK SLIGRSDLNGKRSKIITIENISQPRGIAVHPMAH RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRV SEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHICLDIDECQLGYHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKOMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEOGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 Å 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRPRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM							
ALDHDPVENKIYFAHTALKWIERANMDGSQ RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLNGKRSKIITIENISQPRGIAVHPMAR RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQQCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASIL RLHHRRALDRNKKGYLSRMDLQQIGALAV NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETJDPKKFPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM			1	· · · · · · · · · · · · · · · · · · ·			SPVSWECDCFPGYDLQLDEKSCAASGPQPFL
RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLNGKRSKIITIENISQPRGIAVHPMAR RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVRLQGSMIKPSSLVVVHW LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEEGHATCQCLL GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHIDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MILLLSLWGAHYYRTOKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDQQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASIL RLHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETCDPKKPEPLINSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM		-		ì			LFANSQDIRHMHFDGTDYGTLLSQQMGMVY
SLIGRSDLNGKRSKITTIENISQPRGIAVHPMAR RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTA WCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEIDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM		- 1	i	1			ALDHDPVENKIYFAHTALKWIERANMDGSQ
RLFWTDTGNPRIESSSLQGLGRLVIASSDLIW PSGITIDFLITDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRV SEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPHLREDDHIYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEISS RDVYSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM							RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK
PSGITDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTA WCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHICLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYTEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRITESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM	1		[	ĺ		ĺ	SLIGRSDLNGKRSKITTIENISQPRGIAVHPMAK
PSGITDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTA WCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHICLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYTEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRITESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM			- 1				RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW
RRRLTQNDVGHPFAVAVFEDYVWFSDWAMI SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINI EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIBESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM							PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK
LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITIESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYTEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM	! !	l l					RRRLTQNDVGHPFAVAVFEDYVWFSDWAMP
LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITIESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYTEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM		l	1				SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP
CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM				ŀ			LAKPGADPCLYQNGGCEHICKKRLGTAWCS
KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPFASSKCINT EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM		1	- 1				CREGFMKASDGKTCLALDGHQLLAGGEVDL
VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MILLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM						•	KNQVTPLDILSKTRVSEDNITESQHMLVAEIM
GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM		1	1	}		J	VSDQDDCAPVGCSMYARCISEGEDATCQCLK
BGGYVCRCSEGYQGDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MILLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM	l		l	Ì			
CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MILLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM		į	İ		2		EGGYVCRCSEGYQGDGIHCLDIDECOLGVHS
STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM			l		ļ		CGENASCTNTEGGYTCMCAGRLSEPGLICPD
HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM	1		ł			· .	
DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM				1	.]		
MILLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM				İ	l		
RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ  967 2317 A 8210 3 601 SSAMGSRSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM		. 1					MLLLLSLWGAHYYRTOKILSKNPKNPVEEGG
LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ  LCGMGTEQGCWIPVSSDKGSCPQVMERSFH  MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL  DPPHQMELTQ    SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL  RLHHRFRALDRNKKGYLSRMDLQQIGALAV  NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP  VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY  DLDRDGKISRHEMLQVLRLMVGVQVTEEQL  ENIADRTVQEADEDGDGAVSFVEFTKSLEKM	1	•	- 1	ľ	ľ	İ	RDVRSRRPADTEDGMSSCPOPWEVVIKEHOD
P67 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM			j	J		l	
967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM			1	l	1	1	
967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM		j	J	J			
967 2317 A 8210 3 601 SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM	į	ì	l		1	}	
RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM	967	2317	$\overline{A}$	8210	3	601	
NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM		~	"	3213	-	001	DI UUDED AL DENEVOTA OD COTO COTO
VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM	- 1	ł	ł	1		1	MDI CODINECENDO CONTROLO MARIA A TORRESTA A
DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM		i	l	- 1	1	į	MEDITION OF THE PROPERTY OF TH
ENIADRTVQEADEDGDGAVSFVEFTKSLEKM			ļ	ŀ		ĺ	
	1	ł	- 1	ľ	I	ĺ	
	- 1	ı	i		ŀ	ł	
DVEHKMSIRILK					<u> </u>		DVEHKMSIRILK

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
968	2318	A	8211	2	409	ISSCPHTAYEGSMSTLSNFTQTLEDVFRRIFIT YMDNWRQNITAEQEALQAKVDAENFYYVIL YLMVMIGMFSFIIVAILVSTVKSKRREHSNDP YHQYIVEDWQEKYKSQILNLEESKATTHENIG AAGFKMSP
969	2319	A	8215		1938	GMPRSRGGRAAPGPPPPPPPPGQAPRWSRWR VPGRLLLLLLPALCCLPGAARAAAAAAGAGN RAAVAVAVARADEAEAPFAGQNWLKSYGY LLPYDSRASALHSAKALQSAVSTMQQFYGIP VTGVLDQTTIEWMKKPRCGVPDHPHLSRRRR NKRYALTGQKWRQKHITYSIHNYTPKVGELD TRKAIRQAFDVWQKVTPLTFEEVPYHEIKSDR KEADIMIFFASGFHGDSSPFDGEGGFLAHAYF PGPGIGGDTHFDSDEPWTLGNANHDGNDLFL VAVHELGHALGLEHSSDPSAIMAPFYQYMET HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLPTL PVRIHSPSERKHERQPRPPRPPLGDRPSTPGT KPNICDGNFNTVALFRGEMFVFKDRWFWRL RNNRVQEGYPMQIEQFWKGLPARIDAAYER ADGRFVFFKGDKYWVFKEVTVEPGYPHSLG ELGSCLPREGIDTALRWEPVGKTYFFKGERY WRYSEERRATDPGYPKPITVWKGIPQAPQGA FISKEGYYTYFYKGRDYWKFDNQKLSVEPGY PRNILRDWMGCNQKEVERKERRLPQDDVDI MVTINDVPGSVNAVAVVIPCILSLCILVLVYTI FQFKNKTGPQPVTYYKRPVQEWV
970	2320	A	8216	1235	2223	SRLSLQFYVSFRRTGLFTCKLIVEIFFRNYMN DSLRTNVFVRFQPETIACACIYLAARALQIPLP TRPHWFLLFGTTEEE!QE!C!ETLRLYTRKKPN YELLEKEVEKRKVALQEAKLKAKGLNPDGTP ALSTLGGFSPASKPSSPREVKAEEKSPISINVK TVKKEPEDRQQASKSPYNGVRKDSKRSRNSR SASRSRSRTRSRSRSHTPRRHYNNRRSRSGTY SSRSRSRSHSESPRRHHNHGSPHLKAKHTR DDLKSSNRHGHKRKKSRSRSQSKSRDHSDAA KKHRHERGHHRDRRERSRSFERSHKSKHHGG SRSGHGRHRR
971	2321	A	8217	3	3274	DCRLQAAMPTNFTVVPVEAHADGGGDETAE RTEAPGTPEGPEPERPSPGDGNPRENSPFLNN VEVEQESFFEGKNMALFEEEMDSNPMVSSLL NKLANYTNLSQGVVEHEEDEESRREAKAPR MGTFIGVYLPCLQNILGVILFLRLTWIVGVAG VLESFLIVAMCCTCTMLTAISMSALATNGVVP AGGSYYMISRSLGPEFGGAVGLCFYLGTTFA GAMYILGTIEIFLTYISPGAAIFQAEAAGGEAA AMLHNMRVYGTCTLVLMALVVFVGVKYVN KLALVFLACVVLSILAIYAGVIKSAFDPPDIPV CLLGNRTLSRRSFDACVKAYGIHNNSATSAL WGLFCNGSQPSAACDEYFIQNNVTEIQGIPGA ASGVFLENLWSTYAHAGAFVEKKGVPSVPV AEESRASTLPYVLTDIAASFTLLVGIYFPSVTG IMAGSNRSGDLKDAQKSIPTGTILAIVTTSFIY LSCIVLFGACIEGVVLRDKFGEALQGNLVIGM LAWPSPWVIVIGSFFSTCGAGLQTLTGAPRLL QAIARDGIVPFLQVFGHGKANGEPTWALLLT VLICETGILIASLDSVAPILSMFFLMCYLFVNL ACAVQTLLRTPNWRPRFKFYHWTLSFLGMSL CLALMFICSWYYALSAMLIAGCTYKYIEYRG AEKEWGDGIRGLSLNAARYALLRVEHGPPHT KNWRPQVLVMLNLDAEQAMKHPRLLSFTSQ

		[			.	LVPVKDASRICSLTYLLGSHWNNLVVRSPVL G
974	2324	A	8247	279	468	EYKQWERRFLSCQNRNDLGYGKPRKGGGLL
		ļ		İ	٠.	LSLFGARATSSGGPPLTAGPQREPGARPEPVR SKLPSNL
		ļ	j	Į		GGLGPPGRPYGAIPGQHVTLPRKTSSGSLPPP
		ļ				RGTPPLTPGRLTQDLKLISASQPALPQDGAQT LRRASPHSSGESMAAFPLFPRAGGGSGSGSSS
	1	ł	ĺ	ł	[	RTFPSAPPRASGSHGSLLLPPASSPPPPQVPQR
		-	İ		ł	PSLVAGASGGASPVGFTPRGGLSPPGHSPGPP
		.	ļ			AAQPSPAPPGARGGLGLPEHFLPPPPSSRSPSS SPGQLGQPPGELSLGLATGPLSTPETPPRQPEP
. [	ĺ	[	.		1	GFGHFHKALGGSLSSSDSPLLTPLQPGARSPQ
0	}		ļ		1	PLLPSSSSSPPPGACGSPSAPTPSAGVAATTIA
	l	. [	ľ		j	PPPGSGLGNLGAGQTPRHLKRLQSLIPSALGS ASPASSPSQVDTPSSSSFHIQQLAGFSAPAGLS
	1	I	ļ	}	}	TPLIQAPLQAAAATTSVAIALTHHPRLPAAIFR
		ļ			İ	QIVQHDREMAHCAHRVQAAASATPTPTPVIW
	ł			Ì	ł	LYSLSVDNFNEVLEEYPMMRRAFETVALDRL DRIGKKNSILLHKVQHDLNSGVFNYQENEIIO
	ľ	ĺ	l			TKLADGSYFGEICLLTRGRRTASVRADTYCR
			•			RKLVASMPLFANADPNFVTSMLTKLRFEVFQ PGDYIIREGTIGKKMYFIQHGVVSVLTKGNKE
	- 1	ĺ				YYEHRYQGKMFDEESILGELSEPLREEINFNC
]						WLTMLSMIVGATCYAMFIGHATALIQSLDSS RRQYQEKYKQVEQYMSFHKLPPDTRORIHD
						QYSYALFKAMSHMLCIGYGRQAPVGMSDV
	j		j	. ]		CLQFLVPMLQDFPDDCWVSINNMVNNSWGK
			1	.		TARALRIVRFTKILSLLRLLRLSRLIRYIHQWE EIFHMTYDLASAVVRIVNLIGMMLLLCHWDG
.						YLKSWFMVDFISSIPVDYIFLIVETRIDSEVYK
						LLMVGNLIIIPVGITFFKDENTTPWIVFNVVSD TFFLIDLVLNFRTGIVVEDNTEIILDPQRIKMK
<b>[</b>						VEREQERVKSAGFWIIHPYSDFRFYWDLTML
						CEQPSVDTAIKVEGGAAAGDQILPEAEVRLG QAGFMQRQFGAMLQPGVNKFSLRMFGSOKA
					1	GLAAEPERPGASAQPAASPPPPQQPPQPASAS
				i	ĺ	SGSSHGHLHDSAEERRLIAEGDASPGEDRTPP
						PLPSPSPSAAAGGTESRSSALGAADSEGPARG AGKSSTNGDCRRFRGSLASLGSRGGGSGGTG
					·	AKAWIMDEEEDAEEEGAGGRQDPSRRSIRLR
			/		.020	PRPGPTPAPGAMDKLPPSMRKRLYSLPQQVG
973	2323	A	8237	873	4610	LERKAKSRQVGKEKGKYKEELIEKMQE GCPHAGGKGRVPTGGLTGGRTWSPSAAPRSC
)			ļ.			REKANGTTVHVGIHPSKVVTTRLKLDKDRKKI
	·					VVRGHYKGQQIGKVVQVYRKKYVIYIERVO
1, 7/2	2322	A	8224	701	246	TSRRVTMKFNPFVTSDRSKNRKRHFNAPSHV RRKIMSSPLSKELRQKYNVRSMPIRKDDEVQ
972	2322	A	9224	701	246	ENYMEFLEVLTEGLNRVLLVRGGGREVITIYS
				·		KLNGVVLNKSQDAQLVLLNMPGPPKNRQGD
						HTAAAARTQAPPTPDKVQMTWTREKLIAEK YRSRDTSLSGFKDLFSMKPDQSNVRRMHTAV
ł					}	EQRSQMLKQMQLSKNEQEREAQLIHDRNTAS
	· .					KVWRKCRMRIFTVAQVDDNSIQMKKDLQMF LYHLRISAEVEVVEMVENDISAFTYERTLMM
						ERFGGGHIDVWWIVHDGGMLMLLPFLLRQH
ĺ	ĺ				{	KNFVDTVRDTTAAHQALLVAKNVDSFPQNQ
	]	)	J			NIRSLMSTEKTKGFCQLVVSSSLRDGMSHLIQ SAGLGGLKHNTVLMAWPASWKQEDNPFSW
					i -	LKAGKGLTTVGSVLEGTYLDKHMEAQRAEE
		L .		sequence		nucleotide insertion
				residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
	1	ł		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
uence			914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
cotide seq-	seq- uence		USSN 09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine O=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid converse (A-Alesias C. C.

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- cotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	<b>1</b>	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
[				peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
975	2325	Α	8249	62	1571	LVALKNWKPKGTNIPAPQSPVFGEAVSGVYM
						MTKVLGMAPVLGPRPPQEQVGPLMVKVEEK
						EEKGKYLPSLEMFRQRFRQFGYHDTPGPREA   LSQLRVLCCEWLRPEIHTKEQILELLVLEQFLT
'						ILPQELQAWVQEHCPESAEEAVTLLEDLEREL
						DEPGHQVSTPPNEQKPVWEKISSSGTAKESPS
						SMQPQPLETSHKYESWGPLYIQESGEEQEFAQ
			'			DPRKVRDCRLSTQHEESADEQKGSEAEGLKG DIISVIIANKPEASLERQCVNLENEKGTKPPLQ
						EAGSKKGRESVPTKPTPGERRYICAECGKAFS
1	·					NSSNLTKHRRTHTGEKPYVCTKCGKAFSHSS
						NLTLHYRTHLVDRPYDCKCGKAFGQSSDLLK HQRMHTEEAPYQCKDCGKAFSGKGSLIRHYR
				•		IHTGEKPYQCNECGKSFSQHAGLSSHQRLHT
						GEKPYKCKECGKAFNHSSNFNKHHRIHTGEK
						PYWCHHCGKTFCSKSNLSKHQRVHTGEGEA P
976	2326	Α	8257	298	7086	GNMACWPQLRLLLWKNLTFRRRQTCQLLLE
1						VAWPLFIFLILISVRLSYPPYEQHECHFPNKAM
			l			PSAGTLPWVQGIICNANNPCFRYPTPGEAPGV
					·	VGNFNKSIVARLFSDARRLLLYSQKDTSMKD MRKVLRTLQQIKKSSSNLKLQDFLVDNETFS
		' i	i			GFLYHNLSLPKSTVDKMLRADVILHKVFLOG
		l				YQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP
	1					REKLAAAERVLRSNMDILKPILRTLNSTSPFPS KELAEATKTLLHSLGTLAQELFSMRSWSDMR
[ ]	1		ľ			QEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEG
		ŀ				GGLKIKSLNWYEDNNYKALFGGNGTEEDAE
	ļ	j				TFYDNSTTPYCNDLMKNLESSPLSRIIWKALK PLLVGKILYTPDTPATRQVMAEVNKTFQELA
i						VFHDLEGMWEELSPKIWTFMENSQEMDLVR
			- 1	. 1		MLLDSRDNDHFWEQQLDGLDWTAQDIVAFL
] ]		ŀ	]	·		AKHPEDVQSSNGSVYTWREAFNETNQAIRTIS RFMECVNLNKLEPIATEVWLINKSMELLDER
1	Į.	1	İ			KFWAGIVFTGITPGSIELPHHVKYKIRMGIDN
		[	.			VERTNKIKDGYWDPGPRADPFEDMRYVWGG
1 . 1		Į		j	ļ	FAYLQDVVEQAIIRVLTGTEKKTGVYMQQMP YPCYVDDIFLRVMSRSMPLFMTLAWIYSVAV
l i		- 1	•			IIKGIVYEKEARLKETMRIMGLDNSILWFSWFI
		1	ł			SSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFV
	1	ł	l		ļ	FLSVFAVVTILQCFLISTLFSRANLAAACGGII
				0.0	4	YFTLYLPYVLCVAWQDYVGFTLKIFASLLSP VAFGFGCEYFALFEEQGIGVQWDNLFESPVE
					.	EDGFNLTTSVSMMLFDTFLYGVMTWYIEAVF
	1	1	ł	• 1	ł	PGQYGIPRPWYFPCTKSYWFGEESDEKSHPGS
	į	ŀ	ì			NQKRISEICMEEEPTHLKLGVSIQNLVKVYRD GMKVAVDGLALNFYEGQITSFLGHNGAGKT
			1	]		TTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQ
		- 1	l	ľ	1	NLGVCPQHNVLFDMLTVEEHIWFYARLKGLS
					}	EKHVKAEMEQMALDVGLPSSKLKSKTSQLS GGMQRKLSVALAFVGGSKVVILDEPTAGVDP
]	1				l	YSRRGIWELLLKYRQGRTIILSTHHMDEADVL
[ [	1			1	ĺ	GDRIAIISHGKLCCVGSSLFLKNQLGTGYYLT
.	ŀ	- }			Į	LVKKDVESSLSSCRNSSSTVSYLKKEDSVSQS SSDAGLGSDHESDTLTIDVSAISNLIRKHVSEA
				. 1	Ĭ	RLVEDIGHELTYVLPYEAAKEGAFVELFHRID
	İ	- {	İ		İ	DRLSDLGISSYGISETTLEEIFLKVAEESGVDA
					ļ	ETSDGTLPARRNRRAFGDKQSCLRPFTEDDA
l j					l	ADPNDSDIDPESRETDLLSGMDGKGSYQVKG WKLTQQQFVALLWKRLLIARRSRKGFFAQIV
·						TIPLI YYYI YALLANARALLANARAUTTAYIV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted begiming nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, I=possible nucleotide deletion, \=possible
	·			sequence		nucleotide insertion  LPAVFVCIALVFSLIVPPFGKYPSLELQPWMY NEQYTFVSNDAPEDTGTLELLNALTKDPGFG TRCMEGNPIPDTPCQAGEEEWTTAPVPQTIM DLFQNGNWTMQNPSPACQCSSDKIKKMLPV CPPGAGGLPPPQRKQNTADILQDLTGRNISDY LVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVS NTQALPPSQEVNDATKQMKKHLKLAKDSSA DRFLNSLGRFMTGLDTRNNVKVWFNNKGW
						HAISSFLNVINNAILRANLQKGENPSHYGITAF NHPLNLTKQQLSEVAPMTTSVDVLVSICVIFA MSFVPASFVVFLIQERVSKAKHLQFISGVKPVI YWLSNFVWDMCNYVVPATLVIIIFICFQQKSY VSSTNLPVLALLLLLYGWSITPLMYPASFVFK IPSTAYVVLTSVNLFIGINGSVATFVLELFTDN KLNNINDILKSVFLIFPHFCLGRGLIDMVKNQ AMADALERFGENRFVSPLSWDLVGRNLFAM
						AVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLN DEDEDVRRERQRILDGGGQNDILEIKELTKIY RKRKPAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD
			·		-	GYTTVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV
977	2327	A	8260	3	1567	PGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG
						MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL
978	2328	A -	8261	2	2165	RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE BEAAGTEGDAQEWPGAGSSADQDDEBGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ
						RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE HNLHLLIREFHKQTLSSIMMPHPASAPFGHKR

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
ucnce			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ĺ	ł	residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	l	ł		sequence	ł	/=possible nucleotide deletion, \=possible nucleotide insertion
			<del> </del>	Sequence		MRLSGPQAFDKNEINSLQSSEGLLEKIIKQAK
]	Ì	j	İ		1	HIFLRSRAAATIDSLASRIEDPQIQAHWSNIND
					•	VYESSVKVLITSQGYEQICKSIQLQLNIGVEOI
						RVVHRDGRVITLSYQEQELQDFLLSQMSQHQ
1	Ì		ļ	į .		VHAVQQLAKVMGWQVLSFSNHVGLGPIESIG
	ļ	1		[		NASAITVASPSGDYAISVRNGPESGSKIMVQF
				]	·	PRNQCKDLPKSDVLQDNKWSHLRGPFKEVO
						WNKMEGRNFVYKMELLMSALSPCLL
979	2329	A	8289	2	1053	FVWNPRGGRKRRRQAAVTQAATRASGTPSP
1						RDGTMTQGKLSVANKAPGTEGQQQVHGEKK
]	}		}			EAPAVPSAPPSYEEATSGEGMKAGAFPPAPTA
1						VPLHPSWAYVDPSSSSSYDNGFPTGDHELFTT
						FSWDDQKVRRVFVRKVYTILLIQLLVTLAVV
			1			ALFTFCDPVKDYVQANPGWYWASYAVFFAT
1						YLTLACCSGPRRHFPWNLILLTVFTLSMAYLT GMLSSYYNTTSVLLCLGITALVCLSVTVFSFQ
						TKFDFTSCQGVLFVLLMTLFFSGLILAILLFFQ
						YVPWLHAVYAALGAGVFTLFLALDTQLLMG
[ ]						NRRHSLSPEEYIFGALNIYLDIIYIFTFFLQLFG
					•	TNRE
980	2330	A	8305	59	857	ASQLPDYSISPPSLPPRISFHPSPTLARVAMAEP
]						SEATQSHSISSSSFGAEPSAPGGGGSPGACPAL
						GTKSCSSSCAVHDLIFWRDVKKTGFVFGTTLI
						MLLSLAAFSVISVVSYLILALLSVTISFRIYKSV
1						IQAVQKSEEGHPFKAYLDVDITLSSEAFHNY
						MNAAMVHINRALKLIIRLFLVEDLVDSLKLA
j [						VFMWLMTYVGAVFNGITLLILAELLIFSVPIV YEKYKTQIDHYVGIARDQTKSIVEKIQAKLPG
				[		IAKKKAE
981	2331	A	8308	186	1337	TRMSRHEGVSCDACLKGNFRGRRYKCLICYD
						YDLCASCYESGATTTRHITDHPMQCILTRVD
j				)		FDLYYGGEAFSVEQPQSFTCPYCGKMGYTET
l ·						SLQEHVTSEHAETSTEVICPICAALPGGDPNH
						VTDDFAAHLTLEHRAPRDLDESSGVRHVRR
			· · · · · · · · · · · · · · · · · · ·	Í	j	MFHPGRGLGGPRARRSNMHFTSSSTGGLSSS
				Į	1	QSSYSPSNREAMDPIAELLSQLSGVRRSAGGQ
			. ]	[	Ţ	LNSSGPSASQLQQLQMQLQLERQHAQAARQ
		1				QLETARNATRRTNTSSVTTTITQSTATTNIAN TESSQQTLQNSQFLLTRLNDPKMSETERQSM
				Ì		ESERADRSLFVQELLLSTLVREESSSSDEDDR
				1		GEMADFGAMGCVDIMPLDVALENLNLKESN
				ł	1	KGNEPPPPL
982	2332	A	8315	1	1004	GSTHASADAWAQWFCTEALVMGAPVWYLV
] [				j	· •	AAALLVGFILFLTRSRGRAASAGQEPLHNEEL
1	· [	ĺ	. 1	- 1	ĺ	AGAGRVAQPGPLEPEEPRAGGRPRRRRDLGS
	1			1	]	RLQAQRRAQRVAWAEADENEEEAVILAQEE
· 1	I	ł	ļ	}	i	EGVEKPAETHLSGKIGAKKLRKLEEKQARKA
		1	. 1	j	ŀ	QREAEEAEREERKRLESQREAEWKKEEERLR
		J	·		ł	LEEEQKEEEERKAREEQAQREHEEYLKLKEA
		l			Ī	FVVEEEGVGETMTEEQSQSFLTEFINYIKQSK
.	1	l	ł	ł	l	VVLLEDLASQVGLRTQDTINRIQDLLAEGTIT
	i	1		ļ	I	GVIDDRGKFTYTTPEELAAVANFIRQRGRVSIA
983	2333	A	8320	244	1420	ELAQASNSLIAWGRESPAQAPA RRRWRARGGLVPTLAWAEATGAYVPGRDKP
		.	3320	277	1420	DLPTWKRNFRSALNRKEGLRLAEDRSKDPHD
	ŀ	ł				PHKIYEFVNSGVGDFSQPDTSPDTNGGGSTSD
<u> </u>	ł	1			. ]	TQEDILDELLGNMVLAPLPDPGPPSLAVAPEP
		ì	I	ļ	j	CPQPLRSPSLDNPTPFPNLGPSENPLKRLLVPG
						EEWEFEVTAFYRGRQVFQQTISCPEGLRLVGS
				<u></u>		44

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  EVGDRTLPGWPVTLPDPGMSLTDRGVMSYV RHVLSCLGGGLALWRAGQWLWAQRLGHCH TYWAVSEELLPNSGHGPDGEVPKDKEGGVF DLGPFIVGSLGPPDLITFTEGSGRSPRYALWFC VGESWPQDQPWTKRLVMVKVVPTCLRALVE MARVGGASSLENTVDLHISNSHPLSLTSDQY
984	2334	A	8321	1	1243	KAYLQDLVEGMDFQGPGBS  ANMAPVEHVVADAGAFLRHAALQDIGKNIY TIREVVTEIRDKATRRILAVLPYELRFKEPLPE YVRLVTEFSKKTGDYPSLSATDIQVLALTYQL EAEFVGVSHLKQEPQKVKVSSSIQHPETPLHIS GFHLPYKPKPPQETEKGHSACEPENLEFSSFM FWRNPLPNIDHELQELLIDRGEDVPSEEEEEEE NGFEDRKDDSDDDGGGWITPSNIKQIQQELE QCDVPEDVRVGCLTTDFAMQNVLLQMGLHV LAVNGMLIREARSYILRCHGCFKTTSDMSRV FCSHCGNKTLKKVSVIVSDDGTLHMHFSRNP KVLNPRGLRYSLPTPKGGKYAINPHLTEDQRF PQLRLSQKARQKTNVFAPDYIAGVSPFVENDI SSRSATLQVRDSTLGAGRRRLNPNASRKKFV KKR
985	2335	Α	8322	352	529	RRNNIRQFIMKVCISGQARWLTPVVPVLWET
986	2336	A	8325	89	1172	EAGRSLELKSLRPAWATWGNPISTKINK KMNPTDIADTTLDESIYSNYYLYESIPKPCTKE GIKAFGELFLPPLYSLVFVFGLLGNSVVVLVL FKYKRLRSMIDVYLLNLAISDLLFVFSLPFWG YYAADQWVFGLGLCKMISWMYLVGFYSGIF FVMLMSIDRYLAIVHAVFSLRARTLTYGVITS LATWSVAVFASLPGFLFSTCYTERNHTYCKT KYSLNSTTWKVLSSLEINILGLVIPLGIMLFCY SMIIRTLQHCKNEKKNKAVKMIFAVVVLFLG FWTPYNIVLFLETLVELEVLQDCTFERYLDYA 1QATETLAFVHCCLNPIIYFFLGEKFRKYILQL FKTCRGLFVLCQYCGLLQIYSADTPSSSYTQS TMDHDLHDAL
987	2337	A	8326	3	470	SLSAMRFLAATFLLLALSTAAQAEPVOFKDC GSVDGVIKEVNVSPCPTQPCQLSKGQSYSVN VTFTSNIQSKSSKAVVHGILMGVPVPPPIPEPD GCKSGINCPIQKDKTYSYLNKLPVKSEYPSIK LVVEWQLQDDKNQSLFCWEIPVQIVSHL
988	2338	A		1205	323	VIKMALAARLLPOFLHSRSLPCGAVRLRTPA VAEVRLPSATLCYFCRCRLGLGAALFPRSAR ALAASALPAQGSRWPVLSSPGLPAAFASFPAC PQRSYSTEEKPQQHQKTKMIVLGFSNPINWV RTRIKAFLIWAYFDKEFSITEFSEGAKQAFAH VSKLLSQCKFDLLEELVAKEVLHALKEKVTS LPDNHKNALAANIDEIVFTSTGDISIYYDEKG RKFVNILMCFWYLTSANIPSETLRGASVFQVK LGNQNVETKQLLSASYEFQREFTQGVKPDWT IARIEHSKLLE
989	2339	A	8349	67	185	MSGFIHQLLIQNLFCVYHTRLKTSQGLCLLSL KSLHPMS
990	2340	A	8361	210		ASPILIPMS ASPILIPQGHDSGEREPFSQTPGLMQPFSIPVQ ITLQGSRRRQGRTAFPASGKKRETDYSDGDPL DVHKRLPSSTGEDRAVMLGFAMMGFSVLMF FLLGTTILKPFMLSIQREESTCTAIHTDIMDDW LDCAFTCGVHCHGQGKYPCLQVFVNLSHPG QKALLHYNEEAVQINPKCFYTPKCHQDRNDL LNSALDIKEFFDHKNGTPFSCFYSPASQSEDVI

BGFLSABECVAMQORIGEIVAEMDVPLHCRT EPSTQEEBQLRAQQSGTYFLSGDKREFFEK GVFDEKORIV VPPEKSINLIGHALHAHDPVEK SITHSFRVQTLARSLGLQMPVVVQGWJFKQP HFGGEVSPHQDASELYTEPLGRVLGVWHAVE DATLENGCLWFIPGSHTSGVSRRWVRAPVGS APGTSFLGSEPARDNSLFYPLGRQGALVLHI GEVVHKSKQNLSDRSRQAYTFHLMEASGTT WSPENVLQVTAELPPPQLYT SEPROLUP AM ALSON SKRGAYTFHLMEASGTT WSPENVLQVTAELPPQLYT WSPENVLQVTAELPPQLYT UVGRISLAKEVGGSAVPSSPEEVELNV GGQVFTRISTLSIPHSLLWKMSPSRRDTAN DLAKDSKGRFFDRDGFLFRYLDYLRDRQVV LPDHPFEKGRIKRAESFYGLPDL VKLLTPDEII KQSFDFFCISDFEDASOGSDTRICPPSSLLPAD RKWGFITVGFGSCTILGREQQADAKTRVPR LVCGRISLAKEVGGETLNEGRPDRAPFRYTS RFYLKRHLMGAPASNFLGFWGLORDK HPVNITLQRSVRPDLTSKKAGDLKGKGDA GEVSRRRWLGDFELL VCGRISLAKEVGGETLNEGRPDRAPFRYTS RFYLKRHLMGAPASNFLGFWGLGGGVLCLLHQ SNTSFIKLNNNGFEDIVVDPSSYFEDEKEIQE DMVTTASTLYFAATEKRFFRNVSLIPENWK ENRQYKRPKHENHKHADVIVAPPTLPGBDEP YTKQFTECGEGVHIFTFDLLIGKKQNEYG YFKQFTECGEGVHIFTFDLLIGKKQNEYG PFGKLFVHEWAHLR WGVFDENNEDOFFYRA KSKKEATRCSAGIGSRRWVGGGSCLSKGLI VCLVLDKSGSMGGKDELNEMOAKHFLLQ TYENGSWVGMVHFDSTATIVNKLIQIKSSDER MTLMAGLPYPLGGTSICSGIKTAFQVIGELH SQLDGSEVLLTDGEDATASCIDEVKQSGAI VHRALGRAADEAVERMSKETGGSHYYGDEA QNNGLDAFGALTSGNTDLSQKSLQLESKGLI LNSNAWMDTVILDSTVGKDTFFLITWNSLPP SISLWPSGTIMENTYDATSKMAYLSIFGTA KVGTWAYNLQAKANPSTLITTIVTSRAANSSV PPITVNAKMIKDVNSPPSBMVYAELLQGYVP VLGANVTAFTESQNGFITEVIELLDNGAGADS FKNDGVSSKYTATFENGVSLLARADG NTARLKLRPPLNRAATPGWVVNGEBANPP RPEIDEDTQTILDFPSRTASGGAFVVSQVPSL PLPDQYPPSQTIDLDATVHEDKILTVTARDD NFDVGKVQRYBIRSSSTLUBESRDALQWI TTDLSPKEANSKESPAFKPENISEENATHIFIAI KSBLKSNLTSKVSNLAQVILFPQANFDDDDT TYPTPTPTPTBARSNGVNISTLVSRADGSVVYNOPPSL PLPDQYPRSQTIDLDATVHEDKILTVTARDD NFDVGKVQRYBIRSSSTLUBESRDALQWI TTDLSPKEANSKESPAFKPENISEENATHIFIAI KSBLKSNLTSKVSNLAQVILFPQANFDDDDT TYPTPTPTPTPTBARSGNNISTLVISKISGSVVIV NFILISTII  SSEPRTGREDHOLLILHTERDSRSSQRAVLKIP RQNPGIFWWFLPSRSSHSASHGSRQVSCQG TODELLKMRNTTAELNSLASSANDGAGE TODELLKMRNTTAELNSLASSANDGAGE TODELLKMRNTTAELNSLASSANDGAGG TODELLKMRNTTAELNSLASSANDGAGG	SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acld residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghrtamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isolcucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Ghrtamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LIKKYDQMAIFHCLFWPSLTLLGGALIVGMV RLTQHLSLLCEKYSTVVRDEVGGKVPYIEQH QFKLCIMRRSKGRAEKS
992 2342 A 8370 906 4 MALSGNCSRYYPREQGSAVPNSPERVYELNY GGVVYPRISTLISHISLUKMPSPKRDTAN DLAKDSKGRFFDROGFLFRYILDYLRDRQVV LPDHFPEKGRLKREAEYFQLPDLVKLLTPDEI KQSPDEFCHISDFEDSAQGSDTRICPPSSLLPAD RKWGFITVGYRGSCTLGREGQADAKFRRYPR ILVGRISLAKEVFGETLUSSRDPDRAPERYTS RPYLKFKHLMGAPASNFILGFWGLGQNQDK HFVNIYLQQRSVIRPDLTSKKAGDLKGKGDA QEVSRRRWLGDPEH 993 2343 A 8379 1 2794 MRNQRHRNDTMDFGDSGRGGGGVLCLLPQ SNTSFIKLNNGFEDIVIVDPSVPEDEKIEQIE DMVTTASTYLFEATEKRFFKNVSLIPENWK ENPQYKRRKHENHKHADVVAAPTIPCORDEP YTKOFTEGGEGGEYHFTPDLLLGKKONEYG PPGKLFVHEWAHLRWGYDEYNEDQPFYRA KSKKEATRCSAGISGRNRVYKCQGGSCLSRA CRIDSTTKLYGKDCQFFPDKVQTEKASIMFM QSIDSVVECNEKTHNQEAPSLQNIKCHFRST WEVISNSEDFKNTTEMYTPPPPPYSLLKRQRI VCLVLDKSGSMGGKDRLNRMNQAAKHFLLQ TVENGSWVGMVHFDSTATIVNKLIQIKSSDER MTLMAGLFTYPLGGTSICSGIKYAFQVIGELH SQLDGSEVLLITDGEDNTASSCTGEVKQSGA1 VHFIALGRAADEAVITYDATSKMYLSIPGTA KVGTWAYNLQAKANPETLITIVTSRAANSSV PPITVNAKMIKDVNSFSPMIVYAELGGYVP VLGANVTAFIESQNGHTEVLEILDNGAGADS FKNDGVYSRYFTAYTENGRYSLKVRAHGOA NTARLKLRPPLNRAAYPGWVVNGEIEANPP RPEIDEDTQTTLEDFSRTASGGAFVVSQVESL LPLPDQVPFSQTITLDATVHEDKIILTWTAPGD NFFDVGKVQRYIRISASILDLRDSFDDALQVN TTDLSPKEANSKESFAFKPENISEENATHEIPIAI KSDLKSNLTSKVSNLAQVTLFIPQANPDDIDPT FYTPTPTPTDKSHNSGVNISTLVLSVIGSVVIV NFILSTI  994 2344 A 8385 231 644 INSSPRTGRDHQELNLHTERDSRSQRAVLKIP RQNPGIFYWELLPSRSSISASHGSGRQVSCQG TOPDELLKMRNTFAELKNSLEALSSRMDQAGE	991	2341	A	8369	9	921	SSVVEFSALSVSMACLSPSQLQKFQQDGFLVL EGFLSAEECVAMQQRIGEIVAEMDVPLHCRT EFSTQEEEQLRAQGSTDYFLSSGDKIRFFFEK GVFDEKGNFLVPPEKSINKIGHALHAHDPVFK SITHSFKVQTLARSLGLQMPVVVQSMYIFKQP HFGGEVSPHQDASFLYTEPLGRVLGVWIAVE DATLENGCLWFIPGSHTSGVSRRMVRAPVGS APGTSFLGSEPARDNSLFVPTPVQRGALVLIH GEVVHKSKQNLSDRSRQAYTFHLMEASGTT
SINTSPIKLINNINGFEDIVIVIDPSVPEDEKIEQIE DMYTTASTYLFEATEKRFFFKNVSILIPENWK ENPQYKRPKHENHKHADVIVAPPTLPGRDEP YTKQFTEGCEKGEYHFTPDLLLGKKQNEYG PPGKLFVHEWAHLRWGVFDEYNEDOPFYRA KSKKIEATRCSAGISGRNRVYKCQGGSCLISRA CRIDSTIKLYGKDCQFFPDKVQTEKASIMFM QSIDSVVEFCNEKTHNQEAPSLQNIKCNFRST WEVISNSEDFKNTIPMVTPPPPVFSLLKIRQRI VCLVLJKSGSMGGKDRLNRMNQAAKHFILQ TVENGSWVGMVHFDSTAITVNKLIQIKSSDER NTLMAGLPTYPLGGTSICSGIKYARQVIGEH SQLDGSEVLLTTDGEDNTASSCIDEVKQSGAI VHFIALGRAADEAVEMSKITGGSHFYVSDEA QNNGLIDAFGALTSGNTDLSQKSLQLESKGLT LNSNAWMDIVIIDSTVGKDTFFLITVNSLIPP SISLWDPSGTIMENFTVDATSKMAYLSIPGTA KVGTWAYNLQAKANPETLTITVTSRAANSSV PPPITVNAKMKDVNSFPSPMIVYAEHLQGYVP VLGANVTAFIESQNGHTEVLELLDNGAGADS FKNDGVYSRYFTAYTENGRYSLKVRAHGGA NTARLKLRPPLNRAAYIPGWVYNGEBANPP RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSL PLPDQYPPSQTIDLATVHEDKILTWTAPGD NFFDVGKVQRYIRISASLDLRDSFDDALQVN TTDLSPKEANSKESFAFKPENISEENATHIFIAI KSIDKSNLTSKVSNIAQVTLFIPQANPDDDDPT PTTPTPIPTPDKSHNSGVNISTLVLSVIGSVVIV NFILSTIT  994 2344 A 8385 231 644 INSSPRTGRDHQELNLHTERDSRSQRAVLKIP RQNPGIFYWFLPSRSHSSHGSRGRQVCQG TQDEILKMRNTFAELKNSLEALSSRMDQAEE							MALSGNCSRYYPREQGSAVPNSFPEVVELNV GGQVYFTRHSTLISIPHSLLWKMPSPKRDTAN DLAKDSKGRFFIDRDGFLFRYILDYLRDRQVV LPDHFPEKGRLKREAEYFQLPDLVKILTPDEI KQSPDEFCHSDFEDASQGSDTRICPPSSLLPAD RKWGFITVGYRGSCTLGREGQADAKFRRVPR ILVCGRISLAKEVFGETLNESRDPDRAPERYTS RFYLKFKHLMGAPASNFILGFWGLGQNQDK HPVNIYLQQRSVIRPDLTSKKAGDLKGKGDA QEVSRRRWLGDPEHL
994 2344 A 8385 231 644 INSSPRTGRDHQELNLHTERDSRSQRAVLKIP RQNPGIFYWIFLPSRSHSASHGSRQRQVSCQG TQDEILKMRNTFAELKNSLEALSSRMDQAEE	993	2343	A	8379		2794	MRMQRHKNDTMDFGDSGKRIGGGVLCLLHQ SNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIE DMVTTASTYLFEATEKRFFFKNVSILIPENWK ENPQYKRPKHENHKHADVIVAPPTLPGRDEP YTKQFTECGEKGEYJHFTPDLLLGKKQNEYG PPGKLFVHEWAHLRWGVFDEYNEDQPFYRA KSKKIEATRCSAGISGRNRVYKCQGGSCLSRA CRIDSTTKLYGKDCQFFPDKVQTEKASIMFM QSIDSVVEFCNEKTHNQEAPSLQNIKCNFRST WEVISNSEDFKNTIPMVTPPPPPVFSLLKIRQRI VCLVLDKSGSMGGKDRLNRMNQAAKHFLLQ TVENGSWVGMVHFDSTATIVNKLIQIKSSDER NTLMAGLPTYPLGGTSICSGIKYAFQVIGELH SQLDGSEVLLLTDGEDNTASSCIDEVKQSGAI VHFIALGRAADEAVIEMSKITGGSHFYVSDEA QNNGLIDAFGALTSGNTDLSQKSLQLESKGLT LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPP SISLWDPSGTIMENFTVDATSKMAYLSIPGTA KVGTWAYNLQAKANPETLTITVTSRAANSSV PPITVNAKMNKDVNSFPSPMIVYAEILQGYVP VLGANVTAFIESQNGHTEVLELLDNGAGADS FKNDGVYSRYFTAYTENGRYSLKVRAHGGA NTARLKLRPPLNRAAYIPGWVVNGEIEANPP RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSL PLPDQYPPSQTIDLDATVHEDKIILTWTAPGD NFDVGKVQRYIIRISASILDLRDSFDDALQVN TTDLSPKEANSKESFAFKPENISEENATHIFIAI KSIDKSNLTSKVSNIAQVILFIPQANPDDIDPT PTPTPTPTPIPDKSHNSGVNISTLVLSVIGSVVIV
	994	2344	A	8385	231	644	INSSPRTGRDHQELNLHTERDSRSQRAVLKIP RQNPGIFYWIFLPSRSHSASHGSRQRQVSCQG TQDEILKMRNTFAELKNSLEALSSRMDQAEE
995         2345         A         8390         194         3421         AWRKSSVVPPRGTRRGEKSDQDKSGQKNKR	995	2345	A	8390	194	3421	RIGTQAGVQWRDHGSLQPQPPEFKQCFHLSL PSSWDYRACLS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
996	2346	A	8392	199	3085	DFLSMKQSPALAPEERCRRAGSPKPVLRADD NNMGNGCSQKLATANLLRFLLLVLIPCICALV LLLEILLSYVGTLQKVYFKSNGSEPLVTDGEI QGSDVILTNTTYNQSTVVSTAHPDQHVPAWT TDASLPGDQSHRNTSACMNITHSQCQMLPYH ATLTPLLSVVRNMEMEKFLKFFTYLHRLSCY QHIMLFGCTLAFPECIIDGDDSHGLLPCRSFCE AAKEGCESVLGMVNYSWPDFLRCSQFRNQT ESSNVSRICFSPQQENGKQLLCGRGENFLCAS GICIPGKLQCNGYNDCDDWSDEAHCNCSENL FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVD KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPP CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGQ CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGBADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGR TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHIPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE DISETGYVRPVCLPNPEQWLEPDTYCYTTGW GHMGNKMPFKLQEGEVRISLEHCQSYFDMK TTTTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCPSKVLGPGVYSNVS YFVEWIKRQIYIQTFLLN
			-		3063	KVILSSEMSKTNKSKSGSRSSRSRSASRSRSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSRSPSHN RERNHPRVYQNRDFRGHNRGYRRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWQNYRQAY SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKSSS DRSRRSSSRSSSNHSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTINIDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS SSSQSSHSYKAEEYTEETEEREESTTGFDKSRL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino gold goggest (4 - 41i- C. C.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
þ	1			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1	Į	l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	1	/=possible nucleotide deletion, \=possible
	1	l		sequence		nucleotide insertion
						GTKDFVGPSERGGGRARGTFQFRARGRGWG
1	İ	i	ĺ		1	RGNYSGNNNNNSNNDFQKRNREEEWDPEYT
			]	- 20		PKSKKYYLHDDREGEGSDKWVSRGRGRGAF
	•					PRGRGRFMFRKSSTSPKWAHDKFSGEEGEIE
						DDESGTENREEKDNIOPTTE
997	2347	A	8398	202	552	CPALGGRQDLQGTRLLWAHDSGVGGQKAKS
1		ļ	!		Ì	KQENLESLEATGREEEGGQGPPVTTKGVLLA
		i	l l		}	LLMAGLALQPGTALLCYSCKAQVSNEDCLO
						VENCTQLGEQCWTARIREWGDDSRQA
998	2348	Α	8400	697	301	NPPSACTPGSCDSCSGRGRDLAFDSVWSTNN
		1				MSDPRRPNKVLRYKPPPSECNPALDDPTPDY
						MNLLGMIFSMCGLMLKLKWCAWVAVYCSFI
		1				SFANSRSSEDTKQMMSSFMLSISAVVMSYLQ
1000	00.10	<u> </u>				NPQPMTPPW
999	2349	Α	8401	93	1126	ASASHITSGHLRCFPGSEGVGTMARCFSLVLL
						LTSIWTTRLLVQGSLRAEELSIQVSCRIMGITL
1						VSKKANQQLNFTEAKEACRLLGLSLAGKDQ
1				•		VETALKASFETCSYGWVGDGFVVISRISPNPK
1 !						CGKNGVGVLIWKVPVSRQFAAYCYNSSDTW
ì						TNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYS
1 1						VASPYSTIPAPTTTPPAPASTSIPRRKKLICVTE
1 1						VFMETSIMSTETEPFVENKAAFKNEAAGFGG
1						VPTALLVLALLFFGAAAGLGFCYVKRYVKAF
						PFTNKNQQKEMIETKVVKEEKANDSNPNEES
1000	2350	A	8406	2	777	KKTDKNPEESKSPSKTTMRCLEAEV KERCQFVVKPMLSTVGSFLQDLQNEDKGIKT
[. ]					• • • •	AAIFTADGNMISASTLMDILLMNDFKLVINKI
.			[			AYDVQCPKREKPSNEHTAEMEHMKSLVHRL
						FTILHLEESQKKREHHLLEKIDHLKEQLQPLE
			, ,		ļ	QVKAGIBAHSEAKTSGLLWAGLALLSIQGGA
				1		LAWLTWWVYSWDIMEPVTYFITFANSMVFF
1						AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQQ
						HFDVQQYNKLKEDLAKAKESLKQARHSLCL
						QMQVEELNEKN
1001	2351	Α	8410	1400	264	VGFWERPLRSSRWFRRSLRRWEMLARAARG
	-					TGALLLRGSLLASGRAPRRASSGLPRNTVVLF
			ļ	ļ		VPQQEAWVVERMGRFHRILEPGLNILIPVLDR
		- 1	- 1			IRYVQSLKEIVINVPEQSAVTLDNVTLQIDGV
	ĺ		l			LYLRIMDPYKASYGVEDPEYAVTQLAQTTM
						RSELGKLSLDKVFRERESLNASIVDAINQAAD
	1		ļ		· 1	CWGIRCLRYEIKDIHVPPRVKESMQMQVEAE
	1	ļ	ŀ		i	RRKRATVLESEGTRESAINVAEGKKQAQILAS
	1	-			.	EABKAEQINQAAGEASAVLAKAKAKAEAIRI
		1	1	}		LAAALTQHNGDAAASLTVAEQYVSAFSKLA
	- 1		ļ	ľ	Í	KDSNTILLPSNPGDVTSMVAQAMGVYGALT
		1	[	• ]		KAPVPGTPDSLSSGSSRDVQGTDASLDEELDR
1002	2352	$\overline{\mathbf{A}}$	8421	134	<u></u>	VKMS
1002	ا عدد	^	0421	134	941	NRENLLESRMMDPCSVGVQLRTTNECHKTY
	l			1	ł	YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG
·	l	l				NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL
ľ	į	ļ	1			AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC
		- 1	- 1		J	PKCTQIINGSVDVDTEDRQKRKPESDGRTAK
}	İ	l		}		ALRSLQFTNPGRQTEFAPETGKREKRRLTKN
	J	l				ATAGSDRQVIPAKSKVYDSQGLLIFSGMDLC
	- 1		1	1		DCLDEDCLGCFYACPACGSTKCGAECRCDRK
1003	2353	A	8427	3	1416	WLYEQIEIEGGEIIHNKHAG
			5721	- I	1410	TEWGLSGSCPGCSPLEPGSRGRGAAAWRILR
	l	1			}	CRRLPEPSPFLTQPNLAQSQPPAPVPVTDPSVT MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEIT
						MULUAL POPLATIFICATION OF A STATE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghrtamic Acid, R=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  SLDTENIDEILNNADVALVNFYADWCRFSOM LHPIFEEASDVIKEEFPNENQVVFARVDCDQH SDIAQRYRISKYPTLKLFRNGMMMKREYRGQ RSVKALADYIRQQKSDPIQEIRDLAEITTLDRS KRNIIGYFEQKDSDNYRVFERVANILHDDCAF LSAFGDVSKPERYSGDNIIYKPPGHSAPDMVY LGAMTNFDVTYNWIQDKCVPLVREITFENGE ELTEEGLPFLILFHMKEDTESLEIFQNEVARQL ISEKGTINFLHADCDKFRHPLLHIQKTPADCP
						VIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDL HSGKLHREFHHGPDPTDTAPGEQAQDVASSP PESSFQKLAPSEYRYTLLRDRDEL
1004	2354.	A	8432	910	387	GLSRKLRAGFLPGFCRVSPCGSWVVETLVKM ACAAARSPADQDRFICIYPAYLNNKKTIAEGR RIPISKAVENPTATEIQDVCSAVGLNVFLEKN KMYSREWNRDVQYRGRVRVQLKQEDGSLC LVQFPSRKSVMLYAAEMIPKLKTRTQKTGGA DQSLQQGEGSKKGKGKKKK
1005	2355	A	8453	90	530	QSLQQGGSKKGKKK QSHETKMQSGTHWRVLGLCLLSVGVWGQD GNEEMGGITQTPYKVSISGTTVILTCPQYPGSE ILWQHNDKNIGGDEDDKNIGSDEDHLSLKEF SELEQSGYYVCYPRGSKPEDANFYLYLRARG NPGLQNRYHRLFREDHSKGHSQ
1006	2356	A	8458	3	307	AVQRIRHEMNIFRLTGDLSHLAAIVILLLKIW KTRSCAGISGKSQLLFALVFTTRYLDLFTSFIS LYNTSMKVWYAIHRNVFHLQCTGLWTLNLC QLCIFN
1007	2357	A	8459	43	553	GAGAGGDWAAMDKLKKVLSGQDTEDRSGL . SEVVEASSLSWSTRIKGFIACFAIGILCSLLGT VLLWVPRKGLHLFAVFYTFGNIASIGSTIFLM GPVKQLKRMFEPTRLIATIMVLLCFALTLCSA FWWHNKGLALIFCILQSLALTWYSLSFIPFAR DAVKKCFAVCLA
1008	2358	A	8462	487	150	AQDIRSVHSLGQKSTPVKHFRTLSHLHGLPDP PPHWPPQERSPPSHPCMPSHRPQIPQLSNSGPS DPRWGCVGPSMPTSTCLPGAVEASTTKASLP KCPVDSSLPTPEACFL
1009	2359	A	8465	134	954	ETRVKTSLELLRTQLEPTGTVGNTIMTSQPVP NETIIVLPSNVINFSQAEKPEPTNQGQDSLKKH LHAEIKVIGTIQILCGMMVLSLGIILASASFSPN FTQVTSTLLNSAYPFIGPFFFIISGSLSIATEKRL TKLLVHSSLVGSILSALSALVGFIILSVKQATL NPASLQCELDKNNIPTRSYVSYPYHDSLYTTD CYTAKASLAGTLSLMLICTLLEFCLAVLTAVL RWKQAYSDFPGSVLFLPHSYIGNSGMSSKMT HDCGYEELLTS
1010	2360	A	8468	2	473	KYRYRRPYPVMRKICQVGPAGLAFILNISPVA HRVALCHLAGCQEQAAWYHTLQILFFLVSAY FFSCPVPEKYFPGSCDIVGHGHQIFHAFLSICT LSQLEAILLDYQGRQEIFLQRHGPLSVHMACL SFFFLAACSAATAALLRHKVKARLTKKDS
1011	2361	A	8478	5	409	TELSQLEKAHPPADMGRRKSKRKPPPKKKMT GTLETQFTCPPCNHEKSCDVKMDRARNTGVI SCTVCLEEFQTPITCILGNLGFFQRVGRGLESG PCSSGPLCALVQGQSRPEEQVPPSDFCGVRRC RAGFQCQ
1012	2362	A	8481	2810	1652	RTSTQKWQSVFNDSQEHLERFYCNPENDRM RMKYGGQEFWADLNAMNVYETTEFDQLRR LSTPPSSNVNSIYHTVWKFFCRDHFGWREYPE

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, +possible nucleotide deletion, \=possible nucleotide insertion  SVIRLIEEANSRGLKEVRFMMWNNHYILHNS FFRREIKRRPLFRSCFILLPYLQTLGGVPTQAP PPLEATSSQIICPDGVTSANFYPETWVYMHP SQDFIQVPVSAEDKSYRIIYNLFHKTVPEFKYR ILQLRVQNQFLWEKYKRKEYMNRKMFGR DRIINERHLFHGTSQDVVDGICKHNFDPRVCG KHATMFGQGSYFAKKASYSHNFSKKSSKGV HFMFLAKVLTGRYTMGSHGMRRPPPVNPGS VTSDLYDSCVDNFFEPQIFVIFNDDQSYPYFVI QYEEVSNTVSI
1013	2363	A	8488	2	517	IENCRTRLRQAWHEVCGNKMAAPIPQGFSCL SRFLGWWFRQPVLVTQSAAIVPVRTKKRFTP PIYQPKFKTEKEFMQHARKAGLVIPPEKSDRS IHLACTAGIFDAYVPPEGDARISSLSKEGLIER TERMKKTMASQVSIRRIKDYDANFKIKDFPE KAKDIFIEGSPLY
1014	2364	. A	8501	363	17	YIRTGYVYICIIYAQLMYTYYIRTAYVYICILY AQLMYTYVLYTHSLCIHMYSIRTAYVYICIIY AQIMYTYVFYTHRLCIHMYSIRTDYVYICILY AQIMYTYVFYTHSYMSDE
1015	2365	A	8504	3	2190 453	AQLMYTYVFYTHSYMSDE  NSSEHFSQAPQRLSFYSWYGSARLFRFRVPPD AVLRWLLQVSRESGAACTDAEITVHFRSGA PPVINPLGTSFPDDTAVQPSFQVGVPLSTIPRS NASVNVSHPAPGDWFVAAHLPPSSQKIELKG LAPTCAYVFQPELLVTRVVEISIMEPDVPLPQ TLLSHPSYLKVFVPDYTRELLLELRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC RLLLPSPWDRWLQVTAESLVGPLGTVAFSA VAALTACRPRSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPDQPAEPWACSQKFPCHYQIC KNDREELPAVT
		A				KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTPKNLQTVNVDEN
	2367	A	8513	54		LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSPSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA FFNFAGISVTKELSATTRMVLDSLRTVVIWAL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first arnino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SLALGWEAFHALQILGFLILLIGTALYNGLHR PLLGRLSRGRPLAEESEQERLLGGTRTPINDA S
1018	2368	A	8518	324	694	SPFWTEKRRMEKPLFPLVPLHWFGFGYTALV VSGGIVGYVKTGSVPSLAAGLLFGSLAGLGA YQLYQDPRNVWGFLAATSVTFVGVMGMRS YYYGKFMPVGLIAGASLLMAAKVGVRMLM TSD
1019	2369	A	8526	2	1787	VSAAAVNMEPPDAPAQARGAPRILILAVIL AAHPDAQAEVRLSVPPLVEVMRGKSVII.DCT PTGTHDHYMLEWFLTDRSGARPRLASAEMQ GSELQVTMHDTRGRSPPYQLDSQGRLVLAEA QVGDERDYVCVVRAGAAGTAEAAARLNVF AKPEATEVSPNKGTLSVMEDSAQEIATSNSRN GNPAPKITWYRNGQRLEVPVEMNPEGYMTS RTVREASGLISLTSTLYLRLRKDDRDASFHC AAHYSLPEGRHGRLDSPTFHLTLHYPTEHVQ FWVGSPSTPAGWVREGDTVQLLCRGDGSPSP EYTLFRLQDEQEEVLNVNLEGNLTLEGVTRG QSGTYGCRVEDYDAADDVQLSKTLELRVAY LDPLELSEGKVLSLPLNSRAVVNCSVHGLPTP ALRWTKDSTPLGDGPMLSLSSITTDSNGTYVC EASLPTVPVLSRTQNFTLLVQGSPELKTAEIEP KADGSWREGDEVTLICSARGHPDPKLSWSQL GGSPAEPIPGRQGWVSSSLTLKVTSALSRDGI SCEASNPHGNKRHVFHFGTVSPQTSQAGVAV
1020	2370	A	8530	2	1200	MAVAVSVGLLLLVVAVFYCVRRKGGPCCRQ RREKGAP PRVRLLRPSRSRSCRGLLSTRAPGPSPFRSLHS SPLLPHAMKSPFYRCQNTTSVEKGNSAVMGG VLFSTGLLGNLLALGLLARSGLGWCSRRPLR PLPSVPYMLVCGLTVTDLLGKCLLSPVVLAA YAQNRSLRVLAPALDNSLCQAFAFFMSFFGL SSTLQLLAMALECWLSLGHPFFYRRHITLRLG ALVAPVVSAFSLAFCALPFMGFGKFVQYCPG TWCFIQMVHEEGSLSVLGYSVLYSSLMALLV LATVLCNLGAMRNLYAMHRRLQRHPRSCTR DCAEPRADGREASPQPLEELDHLLLALMTV LFTMCSLPVIYRAYYGAFKDVKEKNRTSEEA EDLRALRFLSVISIVDPWIFIIFRSPVFRIFFHKI FIRPLRYRSRCSNSTNMESSL
1021	2371	A	8536	.1	237	RRGEIDMATEGDVELELETETSGPERPPEKPR KHDSGAADLERVTDYAEEKEIQSSNLETAMS VIGDRRSREQKAKQER
1022	2372	A	8537	94	541	RKERRRRRRMEAVVFVFSLLDCCALIFLSV YFIITLSDLECDYINARSCCSKLNKWVIPELIG HTIVTVLLLMSLHWFIFLLNLPVATWNIYRYI MVPSGNMGVFDPTEIHNRGQLKSHMKEAMI KLGFHLLCFFMYLYSMILALIND
1023	2373	A	8540	26	431	RMMKCPQALLAIFWLLLSWVSSEDKVVQSPL SLVVHEGDTVTLNCSYEVTNFRSLLWYKQEK KAPTFLFMLTSSGIEKKSGRLSSILDKKELSSIL NITATQTGDSAIYLCAVEAQCSLVTCSLYSNS TABALOL
1024	2374	A	8544	1731	743	GVRLRYSPIAVVMVGEAGRDLRRRRAVAVT AEKMAVLAPLIALVYSVPRLSRWLAQPYYLL SALLSAAFLLVRKLPPLCHGLPTQREDGNPCD FDWREVEILMFLSAIVMMKNRRSITVEQHIGN IFMFSKVANTILFFRLDIRMGLLYTTLCIVFLM

No. of   N	SEQ ID	SEQ ID	Met	SEQ	Predicted	I D. 2	
nucleotide sed contespondi periode unco correspondi periode unco correspondi periode unco correspondi periode sed correspondi						Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Decision   Society   Soc			1100				D=Aspartic Acid, E=Glutamic Acid,
1025   2375   A   8546   2194   1707   170		1	1				r=Pnenylalanine, G=Glycine, H=Histidine,
1025							I=Isoleucine, K=Lysine, L=Leucine,
mino seid residue of pegtide sequence		dence					M=Methionine, N=Asparagine, P=Proline,
Pesidue of peptide sequence	uence	J	1	914			Q=Glutamine, R=Arginine, S=Serine,
	1	1		Į.		1	T=Threonine, V=Valine, W=Tryptophan,
				1		sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1025	1		ļ	1			/=possible nucleotide deletion, \=possible
1025   2375   A   8546   2194   1707   TUSPIKY STRYKY STRYKT TOLLT LILE/GGGKAMRRYODKKGRAVSWITSEL VIREPTININELY (PARKEL SKAGDNIPEE) PVAS TPTTY SDOENKEDK   TUSPIK TIMES (LANGER) PROPERTY PEKEN ALPITY SPOENKEDK ALPITY SPOENKEDK   TUSPIK TIMES (LANGER) PROPERTY PEKEN ALPITY SPOENKED   TUSPIK TIMES (LANGER) PROPERTY PEKEN ALPITY SPOENKED   TUSPIK TIMES (LANGER) PROPERTY PEKEN ALPITY SPOENKED   TUSPIK TIMES (LANGER) PROPERTY PEKEN ALPITY SPOENKED   TUSPIK TO THE PROPERTY   TUSPIK SPOENKED   TUSPIK	<u></u>	<u> </u>	<u> </u>		sequence		nucleotide insertion
1025   2375   A   8546   2194   1707   TUSPIKY STRYKY STRYKT TOLLT LILE/GGGKAMRRYODKKGRAVSWITSEL VIREPTININELY (PARKEL SKAGDNIPEE) PVAS TPTTY SDOENKEDK   TUSPIK TIMES (LANGER) PROPERTY PEKEN ALPITY SPOENKEDK ALPITY SPOENKEDK   TUSPIK TIMES (LANGER) PROPERTY PEKEN ALPITY SPOENKED   TUSPIK TIMES (LANGER) PROPERTY PEKEN ALPITY SPOENKED   TUSPIK TIMES (LANGER) PROPERTY PEKEN ALPITY SPOENKED   TUSPIK TIMES (LANGER) PROPERTY PEKEN ALPITY SPOENKED   TUSPIK TO THE PROPERTY   TUSPIK SPOENKED   TUSPIK	ł	1		l			TCKPPLYMGPEYIKYFNDKTIDEELERDKRVT
LNPGKVDVQRYTDVSTRYKVSTSPLTKQLY   LILPQGGKEAMRRQDLKGRAVSWITSEEN     VIREFININELYQRAKKLSKAGDNIPEGPVAS     VIREFININELYQRAKKLSKAGDNIPEGPVAS     VIREFININELYQRAKKLSKAGDNIPEGPVAS     VIREFININELYQRAKKLSKAGDNIPEGPVAS     VIREFININELYQRAKKLSKAGDNIPEGPVAS     VIREFININELYQRAKKLSKAGDNIPEGPVAS     VIREFININELYQRAKKLSKAGDNIPEGPVAS     VIREFININELYQRAKKLSKAGDNIPEGPVAS     VIREFININELYQRAKKLSTVACVICLEEPEYRCKIRS     ALPITLYKPVENKDDDDISLAPTINSDEEBGK     VIREFININGDDDISLAPTINSDEEBGK     VIREFININGDDDISLAPTINSDEEBGK     VIREFININGDDISLAPTINSDEEBGK     VIREFININGDDISLAPTINSDEEBGK     VIREFININGDISLAPTINGDELLYHTINGTUNDA     VIREFINING VIREFINING VIREFINING     VIREFINING VIREFINING VIREFINING     VIREFINING VIREFINING VIREFINING     VIREFINING VIREFINING VIREFINING     VIREFINING VIREFINING VIREFINING     VIREFINING VIREFINING VIREFINING VIREFINING VIREFINING VIREFINING VIREFINING	1	1		I		1	WIVEFFANWSNDCOSFAPIYADLSLKYNCTG
LILEQGGKEAMRRYQDIKLGRAVSWITESER	1	i		i .	-		LNFGKVDVGRVTDVSTRVKVSTSPI TKOI PT
1025   2375   A   8546   2194   1707   TVSPHETMASILKCSTVVCVICLEPEYERGE	1	J	i	1	ì		LILFOGGKEAMERPOIDKKGRAVSWIEGEN
1025   2375   A	1 .	Ì	1	i	}		VIREENI NEI VORAKKI SKAGDNIDERODVAS
1025	1	1	Í	1 .		l	
CRYPYCSVVCFRRHEQCNPETRPVEKKIRS	1025	2375	A	8546	2104	1707	
ALPTKTVKPVENKEDDDGSLADFINSDEEEDR   VSLQNLKNI_GESATLRSLLIPH_RQLMYNI_DQGEDK_AKLMRAYMQEPLFVEFADCCLGIV   EPSQNIESS   PSQNIE				1 00.0	21,54	1,,,,	
1026	1		1	1		Ì	ALDOWNING
DQGEDKAKI.MRAYMQEPLFVEFADCCLGIV	J	ı	j	1		ł	ALPIKIVKPVENKDDDDDSIADFLNSDEEEDR
1026   2376   A   8547   1078   594   VGMELPAVNLKVILLGHWLLTTWGCTVFSGS			l				VSLQNLKNLGESATLRSLLLNPHLRQLMVNL
1026			1			j	DQGEDKAKLMRAYMQEPLFVEFADCCLGIV
1027   2377   A   8557   1   340   DFLGPASPQEGGSESSTMTELETAMGMIDV	1006	0000	<u> </u>				EPSQNEES
1027   2377   A   8557   1   340   DFLEPASQUESTEMICELETAMEMIDIO   SSQURSAYQITIDSAEAPADDFAYPEGRSQDAR   GY   SRYSGSEGSTOTILIKGELKYLMEKELPGFLQ   SGKDKDAYDKILKDLDANGDAQVDFSEFIVF   VAAITSACHKYPEKAKULKULKKELPGFLQ   SGKDKDAYDKILKDLDANGDAQVDFSEFIVF   VAAITSACHKYPEKAKULKULKKELPGFLQ   SGKDKDAYDKILKDLDANGDAQVDFSEFIVF   VAAITSACHKYPEKAKULKKELPGFLQ   SGKDKDAYDKILKDLDANGDAQVDFSEFIVF   VAAITSACHKYPEKAKULKKELPGFLQ   SGKDKDAYDKILKDLDANGDAQVDFSEFIVF   VAAITSACHKYPEKAKULKKELPGFLQ   SGKDKDAYDKILKDLDANGDAQVDFSEFIVF   VAAITSACHKYPEKAKULKELPGFLQ   SGKDKDAYDKILKDLDANGDAQVDFSEFIVF   VAAITSACHKYPEKAKULKELPGFLQ   SGKDKDAYDKILKDLDANGDAQVDFSEFIVF   VAAITSACHKYPEKAKULKELPGFLQ   SGKDKDAYDKILKDLDANGDAQVDFSEFIVF   VAAITSACHKYPEKAKULKELPGFLQ   SGKDKDAYDKILKDLDANGDAQVDFSEFIVF   VAAITSACHKYPEKACH   SGKDKKNL   HODGLAUWTIKDRMQPGOFWFGNMDKFYGLIG   VFVDITYPHEKQQEKKKIL   HODGLAUWTIKDRMQPGOFWILPGG   VFVDITYPHEKQQEKKKIL   HODGLAUWTIKDRMQPGOFWILPGG   VFVDITYPHEKQQEKKKIL   HODGLAUWTIKDRMQPGOFWILPGG   VFVDITYPHEKQQEKKKIL   HODGLAUWTIKDRMQPGOFWILPGG   VFVDITYPHSACHTAPILPILSGIL   TVRTPE   EEKLHRDVFLFSVDNMKLPEMTAPILPICULT   VFRTPE   EEKLHRDVFLFSVDNMKLPEMTAPILPILGIL   VFRTPE   SEKLHRDVFLFSVDNMKLPEMTAPILPILGIL   VFRTPE   SEKLHRDVFLFSVDNMKLPEMTAPILPILGIL   VFRTPE   VFRTPE   VFRTPE   SEKLHRDVFLFSVDNMKLPEMTAPILPILDIL   VFRTPE   VF	1026	2376	A	8547	1078	594	VGMELPAVNLKVILLGHWLLTTWGCIVFSGS
	1	1	l			1	YAWANFTILALGVWAVAQRDSIDAISMFLGG
1027   2377   A   8557   1   340   DFLGPASPQEGGSESSTMTELETAMGMIDV   FSRYSGSEGSTGYTLTKGELKVLMEKELPGFLQ   SGKDKDAVDKILKDLDANGDAQVDFSEFIVF   VAAITSACHKYPEKAG   VAAITS	]				'		LLATIFLDIVHISIFYPRVSLTDTGRFGVGMAIL
SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR	1		ļ	[ ]			SLLLKPLSCCFVYHMYRERGGELLVHTGFIG
1027   2377   A   8557   1   340   DFLGPASPQEEGGSESSTMTELETAMGMIDV   FSRYSGSEGSTQTLTKGELKVLMEKELPGFLQ   SGKDKDAVDKLLKDLDANGDAQVDFSEFIVF   VAAITSACHKYFEKAGILK     1028   2378   A   8569   20   963   KMAATLGFLGSWQQWGRGLSARDGSRRLL   LLLGSGQGPQQVQAGQTFEYLKREHSLSKP   YQGEAPR-PCFLRD WELQVHFKIHGQGKKNL   HGDGLAIWYTKDRMQPGPYGFMDKFYGLG   VYVDTYPNEEKQGERVFFYISAMYNIGSLSY   DHERDGRPTELGGCTAIVRNLHYDTFLJRY   VKRHLTIMMDDGKHEWRDCIEVPGVRLPRG   YYFGTSSTTGDLSDNHDVISLKLFELTVERTPE   EEKLHRDVFLPSVDNMKLPSMTAPLPPLSGL   ALFLIVFFSLVFSVFAIVIGIIL-YNK WQEQSRK   RFY     1029   2379   A   8572   1   578   AAAASHRSRARSRPRVSSGPAPRAQSSAG   RVASGLDSAPLCTMARALCRLPRGLWLLLA   HHLFMTTACQEANYGALLRELCLTQFQVDM   EAVGETLWCDWGTRSYRFLADCTWHMAE   KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR   AVRDPPGSILYPFIVVPTVTLLVTALVVWQS   KRIEGIV   DSSTYVKGGSESRHLCLIPDLKGKARTREASSG   SRTCGRTISLCTSAKSSWTYRSGRLSWQSKG   THLTITQALRQPLHRAPLLPQQLCWSPRPLEK   NKAMGRFLLLPLLLLQPPAFLQPGGSTGSGP   SYLYGVTQFKHLSAMGSVEIPFSTYPWELL   ALVPNVRISWRRGHFHGQSFYSTRPSIHKDY   VNRLFLNWTEGGESFLRISNLRKEDQSVYF   CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT   TTWRPSSTTTLAGLRVTESKCHSESWHLSLDT   ALVPNVRISWRRGHFHGQSFYSTRPPSIHKDY   VNRLFLNWTEGGESFLRISNLRKEDQSVYF   CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT   TTWRPSSTTTLAGLRVTESKCHSESWHLSLDT   ALVPNVRISWRRGHFHGQSFYSTRPPSIHKDY   VNRLFLNWTEGGESFLRISNLRKEDQSVYF   CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT   TTWRPSSTTTLAGLRVTESKCHSESWHLSLDT   ARVALAVAVLKTLQILCLLILL WRRRKG   SRAPSSDF   RTRPTICHGLECTVLLLLTGQL   AFDDPQESCAMMWQKYAGSRRSMPLGARL   FHGVFYAGGFAIVYYLJQKFFISRALYYKLAV   EQLQSPPAGRALQPFINHYLKLDRENPYDI   VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW   UDSPRAQRALQPFINHYLKLDRENPYDI   VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW   UDSPRAQRALQPFINHYLKLDRENPYDI   VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW   UDSPRAQRALQPFINHYLKLDRENPYDI   VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW   UDSPRAQRALQPEVYKLSGENGDEVKKE   RRPPELLPGAEPCEPPVOPRRADMGCSAKAR   RRPPELLPGAEPCEPPVOPRRADMGCSAKAR   RRPPELLPGAEPCEPPVOPRRADMGCSAKAR   RRPPELLPGAEPCEPPVOPRRADMGCSAKAR   RRPPELLPGAEPCEPPVOPRRADMGCSAKAR   RRPPELLPGAEPCEPPVOPRRADMGCSAKAR   RRPPELLPGAEPCEPPVOPRRADMGCSAKAR   RRPPE		]	Ì	[			SSODRSAYOTIDSAEAPADPFAVPEGRSODAR
1028   2378   A   8569   20   963		1		i l			
1028   2378	1027	2377	A	8557	1	340	DFLGPASPOFFGGSESSTMTELETAMGMIDV
1028   2378   A   8569   20   963   KMAATLGPLGSWQQWRRCLSARDGSRRLLL   LLLLGSQGPQQVGAQCTFEYLKREHSLSKP   YQGEAPRPCFLRDWELQVHFKHIGQKKNL   HGDGLALWYTKDRMQPGPVFGMMDKFVGLG   VFVDTYPNEKQQRRVPYSISAMVNIOSLSY   VKRILITIMMDIDGKHEWRDCLEVPGVRLPRG   YYFGTSSITGJD SDNHEDVISLKJEFLTVERTPE   EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL   ALFLIVFFSLVFSVPAIVIGILYNKWQEQSRK   RFY   RVSSTSTIGJD SDNHEDVISLKJEFLTVERTPE   EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL   ALFLIVFFSLVFSVPAIVIGILYNKWQEQSRK   RFY   RVSSTSTIGJD SDNHEDVISLKJEFLTVERTPE   EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL   ALFLIVFFSLVFSVPAIVIGILYNKWQEQSRK   RFY   RVSSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTS		Į.			_	2.0	FSRVSGSEGSTOTI TYGEI YVI MEVEL BOELO
1028   2378   A   8569   20   963   KMAATLGPLGSWQQWRRCLSARDGSRRLL   LLLIGSGQGPQVGAAQTTEYLKREHSLSKP	1		l				SCRURDA ADALI RDI DANCOA OMDESEBUE
1028		İ	1				
LILLIGSGGPQQVGAGQITEYJLKREHSLSKP YQGEAPRPCFLRDWELQVHFKHIGQGKKNI. HGDGLAIWYTKDRMQPGPYGNMDKFVGJLQ VFVDTYPNEEKQQERVFFYJSAMYNNGSLSY DHERDGRFTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDICKHEWRDDICKPVGYKLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVPLPSVDTAMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY  1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGI WILLIA HHLPMTTACQEANYGALRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPFGSILYPFIVVFITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THILTTQALROPLHRAPLLPQGLCWSPRPLEK NKAMGRPLLPLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPPSTYYPWEL AIVPNYEISWRHFHGQSFYSTRPPSHIKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRYBLDTRRSGRQJQSIKGTKLTITQAVTTT TTVRPSSTTTAGLRVTSKGHSSWHISLDT AIRVALAVAVLKTVILGLICLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHAICSVVLLILTGQL AFDDFQESCAMMWQKYAGSRSMPLGARIL FHGYFYAGGFAIVYYLIQKFHSRALYYKLAV EQUQSHPEAQRALGPPLNHYLKLIDREHYDI VDAKLKIPVSGSKSEGLLYVHSSRGGFPGRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE	1028	2378	<u> </u>	8560	20	062	
1029 2379 A 8572 1 578 AAAASHRSRARSPRRVSSGPAPRRAQSSKR RYYGELWCOLGVPVDTYPNEEKQQERVPYISAMYNNGSLSY ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSK RYYGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RYY RFY  1029 2379 A 8572 1 578 AAAASHRSRARSPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRGLWLLA HHLPMTTAQQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVNDPPGSLLYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRFLLLPLLLLLLDPPHAPPLPGGSTGSGP SYLYGVTGYRLSASMGGSVEPPSFYPWEL AIVPNVRISWRRGHFHQQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLITTQAVTTT TTWRPSSTTTLAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGILCLLLLWWRRKG SRAPSSDP  1031 2381 A 8580 905 340 RRTAGTYPCFPKPGRTRHALCSVVLLLTGQL AFDDFQESCAMMVQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPBAQEALGPPLNIHYLKLIDREMFVDI VDAKLKIPVSGSKSEGILTVHSSRGGFFQRW HLDEVFLELKDGQQUPVFKLSGENGDEVKKE	1.020	2070	^	3305	20	903	KMAATLGPLGSWQQWRRCLSARDGSRRLLL
HGDGLAIWYTKDRMQFGPVFGNMDKFVGLG VFVDTYPNEEKQQERVFPYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSTTGDLSDNHDVISLKLFELTVEETTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVPAIVIGIILYNKWQEQSRK RFY  1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTHARALCRLPRGG WLLLA HHLFMTTAQEANYGALIRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVADPPGSILYPTIVVPTIVLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTTTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRFILLIPLLILLQPPAFLQPGGSTGSGP SYLVGVTQFHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPSHKDY VNRLFINWTEGGESGFLRISNLKEDQSVYF CRVELDTRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTLAGLRYTESKGHSESWHLSLDT AIRVALAVAVLKTVILGILCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLILTGQL AFDDFQESCAMMVQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPBAQEALGPPLNIHYLKLDREMFYDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQUPVFKLSGENGDEVKKE			l	] '			LLLLGSGQGPQQVGAGQTFEYLKREHSLSKP
VFVDTYPNEEKQQERVPFYISAMYNNIGSISY DHERDGRPTELGGCTAIVRNIHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSTTGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPISGL ALFELVFSLVFSVFAIVIGIILYNKWQEQSRK RFY  1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWGSIKG THLTTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGSTGSGP SYLYGVTQFKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSHKDY VNRLFLNWTISGRESGFLRISNLRKEDQSVYF CRVELDTRSGRQQLQSIKGTKLTTQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGTYPCFPKPGRTRHALCSVVLLLLTGQL AFDDPQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLJQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDREMFVDI VDAKLKIPVSGKSEGILJVHSSRGGFPGRW HLDEVFLELKDGQQQPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGGAFCEPRVGPRRADMGCSAKAR	ŀ						YQGEAPRPCFLRDWELQVHFKIHGQGKKNL
DHERDGRPTELGGCTAIVRNLHYDTFLVIRY			1				HODGLAIWYTKDRMQPGPVFGNMDKFVGLG
VKRHLTIMMDIDGKHEWRDCIEVPGYRLPRG YYFGTSSITGDL SDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY  1029 2379 A 8572 1 578 AAAASHRSRARSRPRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLPMTTACQEANYGALLRELCLTOFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPQQLCWSPRPLEK NKAMGRFLLPLLLLLLDPPAFLOPGGSTGSOP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHGQSFYSTRPPSHKDY VNRLFLLAULLDPFAFLOPGGSTGSOP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHGQSFYSTRPPSHKDY VNRLFLAULLDPFAFLOPGGSTGSOP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHGQSFYSTRPFSHKDY VNRLFLAULLDPFAFLOPGGSTGSOP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHGQSFYSTRPFSHKDY VNRLFLAULLPFAFLEDGQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDPCESCAMWQKYAGSRRSMPLGARIL FHGYFYAGGFAIVYYLQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNHYLKLIDRENFVDI VDAKLKIPVSGSKSEGILYVHISSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGFRRADMGCSAKKAR		l .	ł	ł 1			VFVDTYPNEEKQQERVFPYISAMVNNGSLSY
1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSGPAPRRAÇSSAG RVASGLDSAPLCTMARALCRLPRIGULLLA HHLIPMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVADPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLITTQALRQFLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQFKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGGESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTITAGIRVTESKGHESSWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGTYPCFFKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL PHGVFYAGGFAIVYYLQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGKSEGILTVHSSRGGFFQRW HLDEVFLEIKDGQQPVFKLSGENGDEVKKE RRFRILLPGAEPCEPRVGFRRADMGCSAKAR	1	1		{		į.	DHERDGRPTELGGCTAIVRNLHYDTFLVIRY
1029 2379 A 8572 1 578 AAAASHRSRARSRPRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRI.PRRGI.WILLA HHIPMTTACQEANYGALLRELCLTOFQVDM EAVGETI.WCDWGRTIRSYREI.ADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSII.YPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHI.CLIPDI.KGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRI.SWQSIKG THILTITQALRQPLHRAPIL.PGQI.CWSPRPLEK NKAMGRPLI.PLILLLQPPAFLQPGGSTGSGP SYLYGVTQPKHI.SASMGGSVEIPFSFYYPWEI. AIVPNVRISWRGHFHGQSFYSTRPSIHKDY VNRI.FINWTEGQESGFIRISNI.RKEDQSVYF CRVELDTRRSGRQQI.QSIKGTKLTITQAVTTT TTWRPSSTTTIAGI.RVTESKGHSESWHI.SLDT AIRVALAVAVI.KTVII.GLI.CLI.LLI.WWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLI.LLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARII. FHGVFYAGGFAIVYYI.QKFHSRAI.YYKLAV EQLQSHPEAQBALGPPLNIHYI.KLIDRENFVDI VDAKI.KIPVSGSKSEGIL.YVHSSRGGPFQRW HLDEVFIELK.DGQQIPVFKLSGENGDEVKKE	İ			1		7	VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
1029 2379 A 8572 1 578 AAAASHRSRARSRPRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRI.PRRGI.WILLA HHIPMTTACQEANYGALLRELCLTOFQVDM EAVGETI.WCDWGRTIRSYREI.ADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSII.YPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHI.CLIPDI.KGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRI.SWQSIKG THILTITQALRQPLHRAPIL.PGQI.CWSPRPLEK NKAMGRPLI.PLILLLQPPAFLQPGGSTGSGP SYLYGVTQPKHI.SASMGGSVEIPFSFYYPWEI. AIVPNVRISWRGHFHGQSFYSTRPSIHKDY VNRI.FINWTEGQESGFIRISNI.RKEDQSVYF CRVELDTRRSGRQQI.QSIKGTKLTITQAVTTT TTWRPSSTTTIAGI.RVTESKGHSESWHI.SLDT AIRVALAVAVI.KTVII.GLI.CLI.LLI.WWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLI.LLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARII. FHGVFYAGGFAIVYYI.QKFHSRAI.YYKLAV EQLQSHPEAQBALGPPLNIHYI.KLIDRENFVDI VDAKI.KIPVSGSKSEGIL.YVHSSRGGPFQRW HLDEVFIELK.DGQQIPVFKLSGENGDEVKKE	l						YYFGTSSITGDLSDNHDVISLKLFELTVERTPE
1029 2379 A 8572 I 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTOFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRFLLPLLLLLLQPPAFLOPGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTVRSGRTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGITPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL PHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRILPGAEPCEPRVGPRRADMGCSAKAR	ľ						EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL
1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRGLWLLLA HHLFMTTACQEANYGALLRELCLTOFQVDM EAVGETL WCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFPDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNHYLKLDRENFVDI VDAKLKIPVSGSKSGGLLYVHSSRGGFPGRW HLDEVFLELKDGQQIPVFKLSGENGGEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEFCEPRVGFRRADMGCSAKAR		1 1		i 1			ALFLIVFFSLVFSVFAIVIGIILYNKWOEOSRK
RVASGLDSAPI.CTMARALCRIPRRGI WILLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLITTQALRQPILHRAPLLPGQLCWSPRPLEK NKAMGRPILLPLLILLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHGQSFYSTRPPSIHKDY VNRLFINWTEGQESGFIRSNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTTTQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGTYPFFFKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLQKFHSRALYYKLAV EQLQSHPEAQBALGPPLNIHYLKLDRENFVDI VDAKLKIPVSGSKSEGLLVYHSSRGGFPQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE		L		<u> </u>	ļ		RFY
RVASGLDSAPLCTMARALCRLPRRGLWILLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPILLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGFFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1029	2379	A	8572	1	578	AAAASHRSRARSRPRRVSSGPAPRRAOSSAG
HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETL WCDWGRTTRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLITTQALRQPHRAPLLPQQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHGQSFYSTRPPSHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLILTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1				ļ	ſ	RVASGLDSAPLCTMARALCRLPRRGLWILLA
EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQBALGPPLNHIYLKLIDRENFVDI VDAKLKIPVSGSKSEGILYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGFRRADMGCSAKAR				]	ļ		HHLFMTTACOBANYGALLRELCLTOFOVOM
Stage	1			1	[		EAVGETLWCDWGRTTRSVRFI ADCTWLINAAE
AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLITTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPFFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					1		KI GCEWPNAEVDREEL AVEIGD VED CODICOD
1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPILLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTITAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE				į l			AVRDPPGSII VPRIVVPITVII I VTAI VAIVO
1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTTTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTITAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLILLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE							KELKCIA ATTIATEDATUTA AMÓS
SST CGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTTQALRQPLHRAPILIPGQLCWSPRPLEK NKAMGRPLLIPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1030	2380	Δ	8574	1352	377	
THLTTTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLILLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR		~00	^	55/4	1336	3/2	CDTCCDDTGI CTC AVGCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
NKAMGRPILLPLLLLLOPPAFLOPGGSTGSGP SYLYGVTOPKHI.SASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFINWTEGQESGFRISNLRKEDQSVYTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGILCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGILYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR						. [	SKILUKKISLUISAKSSWIYRSGRLSWQSIKG
SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLCLLLLWWRRRKG SRAPSSDF RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	[				ł		IHLIIIQALKQPLHRAPLLPGQLCWSPRPLEK
SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLCLLLLWWRRRKG SRAPSSDF RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					•		NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP
AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	] ]			i i	ì	1	SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL
VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR		٠,			Į.		AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY
CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR				1	1		VNRLFLNWTEGQESGFLRISNLRKEDOSVYF
1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					ĺ		CRVELDTRRSGROOLOSIKGTKLTITOAVTTT
AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR						1	TTWRPSSTTTIAGLRVTESKGHSESWHLSIDT
SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFYDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR				ĺ	[	1	AIRVALAVAVI.KTVII.GII.CIIII WWRDDEG
1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLILLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					- 1		
AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1031	2381	A	8580	905	340	
FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR				3200		JTV ]	
EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					J	i	AFDDTQESCAMMWQKYAGSRRSMPLGARIL
VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR				1	I	ĺ	THUYFYAGUFAIVYYLIQKFHSRALYYKLAV
1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					I	!	EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI
1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					ļ	1	VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW
1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1000						HLDEVFLELKDGQQIPVFKLSGENGDEVKKE
WAAGAI GVAGII CAVI GALAMA ABOT BYO	1032	2382	A	8593	2558	961	RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR
TOTAL CONTRACTOR OF THE PROPERTY OF THE PROPER							WAAGALGVAGLLCAVLGAVMIVMVPSLIKQ

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		ĺ	(	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
Į.	1	ļ	1	peptide		/=possible nucleotide deletion, \=possible
			<u> </u>	sequence		nucleotide insertion
ì				ļ	1	QVLKNVRIDPSSLSFNMWKEIPIPFYLSVYFFD
1	i	ĺ	ľ		i	VMNPSEILKGEKPQVRERGPYVYREFRHKSNI
į.		l				TFNNNDTVSFLEYRTFQFQPSKSHGSESDYIV
ļ						MPNILVLGAAVMMENKPMTLKLIMTLAFTTL
	ł	]	2			GERAFMNRTVGEIMWGYKDPLVNLINKYFP GMFPFKDKFGLFAELNNSDSGLFTGFTGVONI
ļ						SRIHLVDKWNGLSKVDFWHSDQCNMINGTS
						GQMWPPFMTPESSLEFYSPEACRSMKLMYKE
	<b>!</b>	l				SGVFEGIPTYRFVAPKTLFANGSIYPPNEGFCP
						CLESGIONVSTCRFSAPLFLSHPHFLNADPVL
ļ						AEAVTGLHPNQEAHSLFLDIHPVTGIPMNCSV
1			[ ]			KLQLSLYMKSVAGIGQTGKIEPVVLPLLWFA
					•	ESGAMEGETLHTFYTQLVLMPKVMHYAQYV
ļ	•					LLALGCVLLLVPVICQIRSQEKCYLFWSSSKK
1000	22.00					GSKDKEAIQAYSESLMTSAPKGSVLQEAKL
1033	2383	A	8595	595	767	AHLPDTLLLPPHSPTVPTPKSFQCSQKACFSRS
1034	2384	A	8597	640	124	FCLLLSLVSSSLVSLSLCPPLTQA
1034	2364	A	1 6391	640	164	VTTSCIPFAFGLGVRASERLAEIDMPYLLKYQ
		·		İ	'	PMMQTIGQKYCMDPAVIAGVLSRKSPGDKIL
						VNMGDRTSMVQDPGSQAPTSWISESQVFQTT EVLTTRITELQRRFPTWTPDQYLRGGLCAYSG
						GAGYVRSSQDLSCDFCNDVLARAKYLKRHG
1						F
1035	2385	A	8603	936	204	AMASTLEYSPSPLRRLVGPAAGFSRAARADL
Į I						SWDPMAFFTGLWGPFTCVSRVLSHHCFSTTG
1 1				·		SLSAIQKMTRVRVVDNSALGNSPYHRAPRÇI
						HVYKKNGVGKVGDQILLAIKGQKKKALIVG
[ [			İ	· i	·	HCMPGPRMTPRFDSNNVVLIEDNGNPVGTRI
1036	2206	-,	0606			KTPIPTSLRKREGEYSKVLAIAQNFV
1036	2386	A	8606	1	562	PTRAHSFDLCCSPCRRLLGREEAGEEPTSPV
1				i		TQYLQPRSPEECKMFACAKLACTPSLIRAGSR
						VAYRPISASVLSRPEASRTGEGSTVFNGAQNG VSQLIQREFQTSAISRDIDTAAKFIGAGAATVG
					f	VAGSGAGIGTVFGSLIIGYARNPSLKQQLFSY
				,		AILGFALSEAMGLFCLMVAFLILFAM
1037	2387	A	8615	2	2364	SPGPSLPESAESLDGSQEDKPRGSCAEPTFTDT
						GMVAHINNSRLKAKGVGOHDNAONFGNOSF
	ſ			ĺ	. [	EELRAACLRKGELFEDPLFPAEPSSLGFKDLG
				ŀ	1	PNSKNVQNISWQRPKDIINNPLFIMDGISPTDI
		- 1				CQGILGDCWLLAAIGSLTTCPKLLYRVVPRG
i i	j	ļ				QSFKKNYAGIFHFQIWQFGQWVNVVVDDRL
	1				1	PTKNDKLVFVHSTERSEFWSALLEKAYAKLS
	į	Į			l	GSYEALSGGSTMEGLEDFTGGVAQSFQLQRP
	ł	ł	·	ł	ł	PONLLRLLRKAVERSSLMGCSIEVTSDSELES
]				ł		MTDKMLVRGHAYSVTGLQDVHYRGKMETLI
	i	- 1	i	1	ł	RVRNPWGRIEWNGAWSDSAREWEEVASDIQ MQLLHKTEDGEFWMSYODFLNNFTLLEICNL
	[	1	ĺ	í	. 1	TPDTLSGDYKSYWHTTFYEGSWRTGSSAGGC
			ļ	l	.	RNHPGTFWTNPQFKISLPEGDDPEDDAEGNV
	į		1	l	ļ	VVCTCLVALMQKNWRHARQQGAQLOTIGFV
ŀ	ļ	ļ		·	į	LYAVPKEFQNIQDVHLKKEFFTKYODHGPSEI
·	ſ	İ	1	ſ	ĺ	FTNSREVSSQLRLPPGEYIIIPSTFEPHRDADFL
	Į.	j	- 1		I	LRVFTEKHSESWELDEVNYAEQLOEEKVSED
	ł	ŀ	Ì	!	ļ	DMDQDFLHLFKIVAGEGKEIGVYELQRLLNR
İ	j	1	ŀ	ļ		MAIKFKSFKTKGFGLDACRCMINLMDKDGSG
				ì	l	KLGLLEFKILWKKLKKWMDIFRECDQDHSGT
	ſ	1	Ì	ľ		LNSYEMRLVIEKAGIKLNNKVMQVLVARYA
		1		}		DDDLIIDFDSFISCFLRLKTMFTFFLTMDPKNT
						GHICLSLEQVLGEGWEGICRIAPACPSTPPPPS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	Ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	испсе	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	j	914	ng to first	acid residue	Q=Ghutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		[		peptide		/=possible nucleotide deletion, \=possible
	<u> </u>	L		sequence	<u></u>	nucleotide insertion
						SDVPGPASCPRLFPPWDLLPVSTVAADDHVGI
			l			EAL
1038	2388	Α	8621	3	1494	RSRMARAPLGVLLLLGLLGRGVGKNEELRLY
		ļ	l			HHLFNNYDPGSRPVREPEDTVTISLKVTLTNL
						ISLNEKEETLTTSVWIGIDWQDYRLNYSKDDF
			İ			GGIETLRVPSELVWLPEIVLENNIDGQFGVAY
1			ĺ			DANVLVYEGGSVTWLPPAIYRSVCAVEVTYF
						PFDWQNCSLIFRSQTYNAEEVEFIFAVDNDG
			ļ			KTINKIDIDTEAYTENGEWAIDFCPGVIRRHH
						GGATDGPGETDVIYSLIIRRKPLFYVINIIVPCV
			•			LISGLVLLAYFLPAQAGGQKCTVSINVLLAQT
						VFLFLIAQKIPETSLSVPLLGRFLIFVMVVATLI
						VMNCVIVLNVSQRTPTTHAMSPRLRHVLLEL
]						LPRLLGSPPPPEAPRAASPPRRASSVGLLLRAE
1						ELILKKPRSELVFEGQRHRQGTWTAAFCQSL
						GAAAPEVRCCVDAVNFVAESTRDQEATGEE VSDWVRMGNALDNICFWAALVLFSVGSSLIF
						LGAYFNRVPDLPYAPCIOP
1039	2389	A	8636	1	900	PGRERPGGGGARRRPOHLPALLPSERPDCATL
]				-		QAMENELPVPHTSSSACATSSTSGASSSSGCN
]					:	NSSSGGSGRPTGPQISVYSGIPDRQTVQVIQQ
				1		ALHRQPSTAAQYLQQMYAAQQQHLMLQTA
				l		ALQQQHLSSAQLQSLAAVQQASLVSNRQGST
]				ļ ļ		SGSNVSAQAPAQSSSINLAASPAAAQLLNRA
						QSVNSAAASGIAQQAVLLGNTSSPALTASQA
						QMYLRAQMLIFTPTATVATVQPELGTGSPAR
						PPTPAQVQNLTLRTQQTPAAAASGPTPTQPVL
<u> </u>						PSLALKPTPGGSQPLPTPA
1040	2390	A	8645	98	1388	ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF
1		ţ				HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ
[	1	- 1		ľ		RRDSFSGVKDSNNNSDGKAVAKVKCEARSA
						LTKPKNNHNCKKVSNEEKPKVAIGEECRADE
		ļ				QAFLVALYKYMKERKTPIERIPYLGFKQINLW
		l			•	TMFQAAQKLGGYETTTARRQWKHIYDELGG
.		l		ļ		NPGSTSAATCTRRHYERLILPYERFIKGEEDKP
		- 1				LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP
		- 1	1	ļ		KSKKEKENAPKPQDAAEVSSEQEKEQBTLISQ
	ĺ	I		1		KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD
	1	ļ		İ		PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA
· [	ĺ	ĺ	ĺ	{		PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWOTR
		ļ		ļ		MRRRMTNCPPWOITLPTAP
1041	2391	$\overline{\mathbf{A}}$	8646	113	1492	LLQEMCTKTIPVLWGCFLLWNLYVSSSQTTYP
			30.0		172	GIKARITQRALDYGVQAGMKMIEQMI.KEKK
[	[	ſ		1		LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP
	ļ	. [		]		NISLAFVPGVGIKALTNHGTANISTDWGFESP
	i	ł	Į	ł	ļ	LFVLYNSFAEPMEKPILKNLNEMLCPILASEVK
		l	İ	1		ALNANLSTLEVLTKIDNYTLLDYSLISSPETTE
		ŀ		ł		NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER
	l	- 1		[		SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS
ľ		ľ	ł	ł	l	TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM
•	1	.			1	VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK
	J		ļ			NSTVETIVSMDFVASTSVGLVILGORLVCSLS
	ŀ	İ	1		ſ	LNRFRLALPESNRSNIEVLRFENILSSILHFGVL
	- [	l		1		PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF
	l	- 1	ĺ			LLISTDLKYETSSKQQPSFHVWEGLNLISRQW
		J		[		RGKSAP
1042 ·	2392	A	8672	538	170	ARRIARTRESKAAVSQDNVPALQPGKKKKLR
		1				LGGKKKKFKFFRLPKEFKKQLMYSPSNFKKM
		<del></del>				TOTAL STATE OF STATE

SEOID	SEQ ID	Met	SEQ	Predicted	T 75 - 41 - 4 - 4	
NO: of	NO: of	hod	ID NO:	beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in NO.	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
cotide	seq-		USSN	location		F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence		09/496	correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine,
uence	10000	ľ	914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline,
]		j	114	amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine,
	1			residue of	sequence	T=Threonine, V=Valine, W=Tryptophan,
ł	ľ		1	peptide	Sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	J	1	1	Sequence		/-possible nucleotide deletion, \-possible
<b>—</b>	<del> </del>	<del></del> -	<del> </del>	Sequence	<del> </del>	nucleotide insertion
						TSLAGNTVQCLNKLKYVIYSAQYPAYGNITT
1043	2393	A	8688	359	17	LDMITSTDHVLEQDFWICFTFYSVKERQI
1		<b>  ^</b>	8000	339	1 17	GLKTRAPATPTFQREVLGPAKQDMQRRCPRI
1		ļ.		}		GLMTSLIKPIKRRWRDYKRWKSGGFTGESC
		į	1			HHADTLGDRGGLQGDHSELLQWQKRILRTE
1044	2394	A	8718	292	1490	GEPSPKYISKNIFPICSYITGFL
1000	224	^	0718	1 492	1490	GTVKTSVATPITAGHSCSSGGVLQVKSPATQS
1		l				GFKFTSKMEDFNMESDSFEDFWKGEDLSNYS
		1		l		YSSTLPPFLLDAAPCEPESLEINKYFVVIIYAL
]		ļ				VFLLSLLGNSLVMLVILYSRVGRSVTDVYLL
I		<b>[</b>			1	NLALADLLFALTLPIWAASKVNGWIFGTFLC
1	İ	1				KVVSLLKEVNFYSGILLLACISVDRYLAIVHA
	]	,		·		TRTLTQKRYLVKFICLSIWGLSLLLALPVLLFR
Í	<b>j</b>	ĺ			i	RTVYSSNVSPACYEDMGNNTANWRMLLRIL
Ì						PQSFGFTVPLLIMLFCYGFTLRTLFKAHMGQK
						HRAMRVIFAVVLIFLLCWLPYNLVLLADTLM
						RTQVIQETCERRNHIDRALDATEILGILHSCLN
1	ĺ					PLIYAFIGQKFRHGLLKILAIHGLISKDSLPKDS
1045	2395	A	8724	254	3104	RPSFVGSSSGHTSTTL
1043	2393	A	0/24	254	3184	FRANLAITVANRRGAQGGKMHTCCPPVTLEQ
						DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY
[	1					YGEICDNACPCEEKDGILTVSCENRGIISLSEIS
<b>!</b> .						PPRFPIYHLLLSGNLLNRLYPNEFVNYTGASIL
						HLGSNVIQDIETGAFHGLRGLRRLHLNNNKL
1						ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF
1						GKLHLLQVLILNDNLLSSLPNNLFRFVPLTHL
						DLRGNRLKLLPYVGLLQHMDKVVELQLEEN
1 1						PWNCSCELISLKDWLDSISYSALVGDVVCETP
						PRLHGRDLDEVSKQELCPRRLISDYEMRPQTP
1 1	' 1	i	1	1		LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP
						KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA
i ,				1	ļ	YQTKSPVPLECPTACSCNLQISDLGLNVNCOE
1 1	1	- 1		İ	}	RKIESIAELQPKPYNPKKMYLTENYIAVVRRT
<b>!</b>						DLLEATGLDLLHLGNNRISMIQDRAFGDLTN
				ł		LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY
1 1	Į.			i i	ļ	NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS
ĺ	1			,		GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS
			ŀ	į		LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG
	•	j		J		VLVDEVICKAPKKFAETDMRSIKSELLCPDYS
	ļ	- 1	i	i	[	DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA
	ł		j	ļ		PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA
	ŀ	ļ	l		ļ	AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN
		į	1		İ	MQYSVYGGGGTGGHPHAHVHHRGPALPK
	}		ļ	j	J	VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN
		Į	. ]	{		SVEDYKDLHELKVTYSSNHHLQQQQQPPPPPP
		f	F	l	]	QQPQQPPPQLQLQPGEEERRESHHLRSPAYS
		ł	l	ŀ	l	VSTIEPREDLLSPVQDADRFYRGILEPDKHCST
			. [	[	ĺ	TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH
		ļ		ł	1	QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE
						YLELKAKLNVEPDYLEVLEKQTTFSQF
1046	2396	A	8736	28	452	SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT
1	1			J		AMAGALVRKAADYVRSKDFRDYLMSTHFW
		ł	J	. 1	ı	GPVANWGLPIAAINDMKKSPEIISGRMTFALC
			1		ļ	CYSLTFMRFAYKVQPRNWLLFACHATNEVA
	- I	į				QLIQGGRLIKHEMTKTASA
1047	2397	A	8741	673	924	ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP
		- 1	ļ	-		AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK
	Į.	1	ł		ĺ	PPTTKLLHSSPLWNFFAQQL
1048	2398	A	8747	3	5054	PEVTKPSLSQPTAASPIGSSPSPPVNGGNNAKR
				<del>-</del>	<u>                              </u>	XXANNOONYTICOSTICOSTSTEVNOONNAKK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hođ	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	nence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
neuce			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ł				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1			peptide		/-possible nucleotide deletion, \-possible
				sequence		nucleotide insertion
ļ						VAVPNGQPPSAARYMPREVPPRFRCQQDHK .
1						VLLKRGQPPPPSCMLLGGGAGPPPCTAPGAN
						PNNAQVTGALLQSESGTAPDSTLGGAAASNY
ļ						ANSTWGSGASSNNGTSPNPIHIWDKVIVDGS
						DMEEWPCIASKDTESSSENTTDNNSASNPGSE
1	·	-				KSTLPGSTTSNKGKGSQCQSASSGNECNLGV
-			į			WKSDPKAKSVQSSNSTTENNNGLGNWRNVS
1						GQDRIGPGSGFSNFNPNSNPSAWPALVQEGTS
1						RKGALETDNSNSSAQVSTVGQTSREQQSKME
1						NAGVNFVVSGREQAQIHNTDGPKNGNTNSL NLSSPNPMENKGMPFGMGLGNTSRSTDAPSQ
4				l	•	STGDRKTGSVGSWGAARGPSGTDTVSGQSNS
1				ļ		GNNGNNGKEREDSWKGASVQKSTGSKNDS
	.			ĺ		WDNNNRSTGGSWNFGPQDSNDNKWGEGNK
		ſ	1			MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN
<b>.</b>			ľ			SSTGAWDNQKGHPLLENQGNAQAPCWGRSS
						SSTGSEVEGOSTGSNHKAGSSDSHNSGRRSY
1 1		Į.				RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG
1		1	ŀ			QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES
			Ì			AATQTKNSGGWGDAPSQSNQMKSGWGELS
						ASTEWKDPKNTGGWNDYKNNNSSNWGGGR
1 1		ĺ	- 1			PDEKTPSSWNENPSKDQGWGGGRQPNQGWS
		ì	}			SGKNGWGEEVDQTKNSNWESSASKPVSGWG
		1				EGGQNEIGTWGNGGNASLASKGGWEDCKRS
1 1	ľ	ł	ł	}		PAWNETGRQPNSWNKQHQQQQPPQQPPPPQ
	i		1	1		PEASGSWGGPPPPPPGNVRPSNSSWSGPQPA
1	ļ	1	- 1			TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP
1 }		}				MTSKSASDSKSMQDGWGESDGPVTGARHPS
1 1	.		,			WEEEEDGGVWNTTGSQGSASSHNSASWGQG
1		j	i	!	•	GKKQMKCSLKGGNNDSWMNPLAKQFSNMG
1		- 1		• 1	-	LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA
i i	ĺ	9 [		ļ	1	MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS
]				1	1	GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN
ł I		1				SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP
ł 1	- 1	ł	- 1	ł	1	GLFNVGPQLSPQQIAMLSQLPQIPQFQLACQL
1 1	- 1	- 1	ļ	l	ļ	LLQQQQQQLLQNQRKISQAVRQQQEQQLA
	i			·	I	RMVSALQQQQQQQQQRQPGMKHSPSHPVGPK
]		ļ		Į.	1	PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF
		1				SSGGMDYGMVGGKEAGTESRFKQWTSMME
	ľ	1	l	l	-	GLPSVATQEANMHKNGAIVAPGKTRGGSPY
		ł	- 1		-	NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK
		1				IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS
}	1		į	. }		NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK
						FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS
•	į	ļ	1		l	RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG
] ]	1	l		1	I	WGTQDSRLASASTWSDGGSVRPSYWLVLHN
'[	ĺ	l	ĺ	Í	1	LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA
		- 1	- 1	ł	Ī	LIRYSTKQEAAKAQTALHMCVLGNTTILAEF
1		j		. 1		ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS
·	ł	· }	ļ	- 1	ļ	LETGQNQSDPVGPALNLFGGSTGLGQWSSSA
<b> </b> [	1	- 1			i	GGSSGADLAGASLWGPPNYSSSLWGVPTVED
						PHRMGSPAPLLPGDLLGGGSDSI
1049	2399	A	8748	200	1387	VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF
		- 1		- 1	. [	LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS
		i		I		PQDCGFKDHQPLTLQALTVELARWTLMLLLS
	ļ	[		I		TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT
	1		1	I	l	ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA
<u> </u>	1					PFALSALLYGANNNLVIYLQRYMDPSTYQVL

SEQ ID   No. of nucleoid   DNO:   pepide   Seq   DNO:   pepide	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	I Andrews St. All Control of the Con
nucleotide seq- uence unice un							Amino acid sequence (A=Alanine C=Cysteine,
Sociation   Soci	nucl-		1				F=Phenylalanine G=Chroine U=Uistidine
Sequence	eotide	1	ļ	1			I=Isoleucine K=I veine I =I eucine
Defice   Popular   Popu	seq-	uence	i	•		to last amino	M=Methionine N=Assarsoine P=Proline
amino acid residue of sequence of peptide residue of sequence peptide sequence of peptide sequence of peptide sequence of sequ	uence	1		914			O=Glutamine, R=Arginine, S=Serine
residue of peptide sequence     Peptide of peptide sequence	İ	}	l	İ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryntonhan
Peptide		1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon
	}		j	}	peptide		/=possible nucleotide deletion, \=possible
MAGAGYAGGIQVYGORILSPPPAAASC	<u> </u>		<u> </u>		sequence		nucleotide insertion
MAGAGYAGGIQVYGORILSPPPAAASC	] .		1				SNLKIGSTAVLYCLCLRHRLSVRQGLALLLL
0RLPLALONLEYTTGVLLINGLHAGGGSGG   GLEEFGSWAALVUSQALNGLHASAYMKH   GSSTRIFVVSCSLVVNAVLSAMIGHAGGGSGG   GLEEFGSWAALVUSQANGLHASAYMKH   GSSTRIFVVSCSLVVNAVLSAMILRUQLTAA   FILATLLIGLAMKLYVGSR   RVLLIHELIPPLA VPAHRCALPGAPANDE   RVLLIHELIPPLA VPAHRCALPGAPANDE   HOPDWILSAHLPREPDGTLSSCLRRAVPQALP   MTHOEERGGRGELEDEPATVTCSQGWEVDH   SEPSSTIATESQWDLVCEQKGINAASTFFFA   GYLVGAVAPGVLSDRFORRRLLLVAVVSTILL   LGLASAASVSYVMPATRITLTGSLAGFTIIV   MPILELEWILDVEHRIVAGVLSSTFFYTGGVML   LGLASAASSYVMPATRITLTGSLAGFTIIV   MPILELEWILDVEHRIVAGVLSSTFFYTGGVML   LALVGYLIEDWRWLLAVTLACHGGILSLWW   VPESARWLLTQGIVKKBAHRYLHGCARINGR   PVCEDSTSQBAVSKVAAGERVYRAFGLLVTJSVRYA   GRRITQAGTILLGTALAFGTISLLVSSDMSSW   TVLAWMGKAFSEAAFTTAYLTESLYPTVUR   QTOMGITALVGRIGGSLARLAALLDGVUN   GRILNYYTQLIGAVELPKLLVYLSVSMSSW   TVLAWMGKAFSEAAFTTAYLTESLYPTVUR   QTOMGITALVGRIGGSLARLAALLDGVUN   CTV. MGRAFSEAAFTTAYLTESLYPTVUR   QTOMGITALVGRIGGSLARLAALLDGVUN   CTV. MGRAFSEAAFTTAYLTESLYPTVUR   QTOMGITALVGRIGGSLARLAALLDGVUN   CTV. MGRAFSEAAFTTAYLTESLYPTVUR   QTOMGITALVGRIGGSLARLAALLDGVUN   CTV. MGRAFSEAAFTTAYLTESLYPTVUR   QTOMGITALVGRIGGSLARLAALLDGVUN   CTV. MGRAFSEAAFTTAYLTESLYPTVUR   QTOMGITALVGRIGGSLARLAALLDGVUN   CTV. MGRAFSEAAFTTAYLTESLYPTVUR   QTOMGITALVGRIGGSLARLAALLDGVUN   CTV. MGRAFSEAAFTTAYLTESLYPTVUR   QTOMGITALVGRIGGSLARLAALLDGVUN   CTV. MGRAFSEAAFTTAYLTSRAQALPET   QTV. MGRAFTAVGRIGGSGGGGGETTQDEVSSHTS   EEDGGVVVEKELENTEQPVGGRISVPERPL   CTV. MGRAFSEAGTDDARSTAN   CTV. MGRAFSEAGTDARSTAN   CT	1			1	ł	1	MAAGACYAAGGLOVPGNTLPSPPPAAAASP
1050   2400   A   8758   3   1660   WYSSMGFEELEGVGFGFFOLRIVALIALD   FFLATILIGIANRILYYGSR   FFLATILIGIANRILYYGSR   FFLATILIGIANRILYYGSR   WYSSMGFEELEGVGFGFFOLRIVALIALD   RVILPHIFILIPII-AAVPAHRCALPGAPANPS   HQDVWLAHLIPEPDOTISSCI-RAYQALP   HTTLGERGSGIELDEPATVPCSQGWETDH   SEPSSTIATESQWUD/VECNGGINASTFFTA   GVLVGAVAPGVLSDRFORRRILLIVAYYSTI   LGLASAASYSVAWAGENTATITICOSIA GETTIV   LGLASAASYSVAWAGENVARTSTIATOSIA GETTIV   LGLASAASYSVAWAGENVARTSTIATOSIA GETTIV   LGLASAASYSVAWAGENVARTSTIATOSIA GETTIV   LGLASAASYSVAWAGENVARTSTIATOSIA GETTIV   WPLELEWILDVEHRIVAGVUSSTFWTGGIML   LALVGYIEDWRWILIAYTI-PCAPGIISIML   LALVGYIEDWRWILIAYTI-PCAPGIISIML   LALVGYIEDWRWILIAYTI-PCAPGIISIML   LALVGYIEDWRWILIAYTI-PSELYPTIVA   GRITTOMOTICATA GERILIVYISSMKSWS   TVLAWMGKAFSGAAFTTAYLFTSELYPTIVA   GRITTOMOTICATA GERILIVYISSMKSWS   TVLAWMGKAFSGAAFTTAYLFTSELYPTIVA   GRITTOMOTICATA GERILIVYISSMKSWS   TVLAWMGKAFSGAAFTTAYLFTSELYPTIVA   GRITTOMOTICATA GERILIVYISSMKSWS   TVLAWMGKAFSGAAFTTAYLFTSELYPTIVA   GRITTOMOTICATA GERILIVYISSMKSWS   TVLAWMGKAFSGAAFTTAYLFTSELYPTIVA   GRITTOMOTICATA GERILIVYISSMKSWS   TVLAWMGKAFSGAAFTTAYLFTSELYPTIVA   GRITTOMOTICATA GERILIPSKILIVYISSMKSWS   TVLAWMGKAFSGAAFTTAYLFTSELYPTIVA   GRITTOMOTICATA GERILIPSKILIVYISSMKSWS   TVLAWMGKAFSGAAFTTAYLFTSELYPTIVA   GRITTOMOTICATA GERILIPSKILIVYISSMKSWS   TVLAWMGKAFSGAAFTTAYLFTSELYPTIVA   GRITTOMOTICATA GERILIPSKILIVYISSMKSWS   TVLAWMGKAFSGAAFTTAYLFTSELYPTIVA   GRITTOMOTICATA GERILIPSKILIVYISSMKSWS GRITTOMOTICATA GRITTOMOTICATA GERILIPSKILIVYISSMKSWS GRITTOMOTICATA GRITT	1		ĺ				MPLHITPLGLLLLILYCLISGLSSVYTELLMKR
OSSITRIPVYSCSLVYNAYLSAVILRIQITAA   FILATILIGIAMRLYYGSR   FILATILIGIAMRLYYGSR   WVSSMGFEELLEQVGGFGFPQLRIVALIAIP   RIVELPHILIPIFILAAVPAHREALPGAPANIPS   HQDVWLEAHLPREPPGTISSCLRFAYPQALP   NTITIGERGSRGELEDPAYTGALP   NTITIGERGSRGELEDPAYTGALP   NTITIGERGSRGELEDPAYTGALP   SEPSSTIATESQWDLVCEQKGLNRAASTFFFA   GVVLYGAVAFGYLSDRERGERILLIVAYVSTLV   LGLASAASVSYVMFAITRITTGALAGFTIIV   MPIELEWIDVEHRIVAGVISSTRWTGGVML   LALVOYLIRDWRWILLAAVTLLCAFGILSUW   VPSARWILITCHYCHYCKEAHRYLLHCARINGR   PVCCDSFSQEAVSKVAAGEKVVRRFSYLDLIF   RTFRLRHSLCCVVVWFGVNFSYYGLISDWS   GLGINYYGTQLIFGAVELPSKLIVVISWRYA   GRRITQAGTILATALAFGTRILIVSSDMKSWS   TVLAVMGKAFSAAFTTAVTERYPTVIL   QTGMGLTALVGRLGGSLAPLAALIDGWWLS   LFLLTYGGGLIAAGTALIPESLTYPTVIL   QTGMGLTALVGRLGGSLAPLAALIDGWWLS   LFLLTYGGGLIAAGTALIPESLYPTVIL   QTGMGLTALVGRLGGSLAPLAALIDGWWLS   LFLLTYGGGLIAAGTALIPESLYPTVIL   QTGMGLTALVGRLGGSLAPLAALIDGWWLS   LFLLTYGGGLIAAGTALIPESLYPTVIL   QTGMSLTAAVGRLGGSLAPLAALIDGWWLS   LFLLTYGGGLIAAGTALIPESLYPTVIL   QTGMSLTAAVGRLGGSLAPLAALIDGWWLS   LFLLTYGGGLIAAGTALIPESLYPTVIL   QTGMSLTAAVGRLGGSLAPLAALIDGWWLS   LFLLTYGGGLIAAGTALIPESCH   LFLLTYGGGLIAAGTALIPESCH   LFLLTYGGGLIAAGTALIPESCH   LFLLTYGGGLIAAGTALIPESCH   LFLLTYGGGLIAAGTALIPESCH   LFLLTYGGGLIAAGTALIPESCH   LFLLTYGGGLIAAGTALIPESCH   LFLLTYGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG		1			1		QRLPLALQNLFLYTFGVLLNLGLHAGGGSGP
1050   2400   A   8758   3   1660				i			GLLEGFSGWAALVVLSQALNGLLMSAVMKH
1050	1		ĺ		1		
RVILPHELIPITLAAVPARIKCALPGAPANPS HQDVWLEAHLPREPDGTLSSCLRFAYPQAIP MTIGERQSRGELEDEPATYPCSQG WEYDH SEPSSTIATESQWILVAYVSTLU LGLASAASVSYVMFAITRILTGSALAGFTIV MPLELEWIDVERITVAGVILVAYVSTLU LGLASAASVSYVMFAITRILTGSALAGFTIV MPLELEWIDVERITVAGVILVAYVSTLU LGLASAASVSYVMFAITRILTGSALAGFTIV MPLELEWIDVERITVAGVILAVTLPCARGILSI WW VPESARWILTQGHVKEAHRYLLHCARILOR PVCEDSFSQEAVSKVAAGERVVRRFSYLDLIS RTPRIRHISLCCVVVWFGVNFSYYGLISLDVS GLGINNYQTOLLFGAVELIFSKILVVIJSVRYA GRRILTQAGTILJGAAGTALLPSSDMKSWS TVLAVMGKAFSEAAFTTAYTELYPTVLR QTGMGITALVGRIGGSLAPLAALLDGVWLS LFRLTYGGGALLAAGTALLPSELYPTVLR QTGMGITALVGRIGGSLAPLAALLDGVWLS LFRLTYGGGALLAAGTALLPSELYPTVLR QTGMGITALVGRIGGSLAPLAALLDGVWLS LFRLTYGGGALLAAGTALLPSELYPTVLS LFRLTYGGGALLAAGTALLPSELYPTVLS TONLANDFILELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPOSLGR HLLHISLDQRSNCAVCGARFTSHATTNSEKLP EVLNMSSLIPTVHBGFSSAGGGALAFSPOVTY AGILLVCNNCAAVSKLIEAQTTSVRKWALRR QNEPLEVELQRIBERETAKSREKTLEXT QREPLEVELQRIBERETAKSREKTLEXT LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSLI LGRADAVFSVNHLLGDDFARNLQRDR EAMRLKRANETPEERQAALTPSTRAKTLKRT LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSLI LGRADAVFSVNHLLGDDFANVAMAYMASSI ASHGCMWYMERLHRFVSVSLKKVFFAVDTA ASAUVFVFTRRSGPSGTASVAAMAYHSGYGA HGSKKHRARAAPPPPFLFDTSGGYSSQFGGY PATGADVAFSVNHLLGDVAGMAGTGGRYPFGRGAA ASLVFVFTRRSGPSGTASVAAMAYHSGYGA HGSKKHRARAAPPPPFLFDTSGGYSSQFGGY PATGADVAFSVNHLLGDVAGMAGTGGRYPFGRGAA ASLVFVFTRRRSGPSGTASVAAMAYHSGYGA HGSKKHRARAAPTVSVSLKKVFFAVDTA AYMKKGLLVFPYTHONWEVQYSRDALARICGUNG RESPEVIGLCASTALVWVMEVIALLICLYL ATVRSDLSTFHLIAYSGYKYVOMILSVLTGL LIPGSDGYYVALAWTSSALMYTYNFRSIKTAAL GPDSMGGPVFRQRIQLYLLIGAAAFQHIMY VTFHLVR LIPGSDGYYVALAWTSSALMYTYNFRSIKTAAL GPDSMGGPVFRQRIQLYLLIGGAAFQFLIMY VTFHLVR VTFHLVR SNIKKTADMDVQQGGFHRRYDVVYTTPW SNIKKTADMDVQQGGFHRRYDVVYTTPW SNIKKTADMDVQQGGFHRRYDVVYTTPW SNIKKTADMDVQQGGFHRRYDKLTVTVEKK VPELINILEKTKVERFFPDLAAKEGCDRBGRR NEKKAQQDEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM RICHARD ARMS CARGERPYNCSMP SSLMKVENMSSNQDGNDSDEFM	1050	2400	_	0750		1660	
HQDVWLEAHILREPDGTLSSCLRFAYPQALIP NTTLGERGSGELEDEPATYCSQGWEYDH SERSSTIATESQWDLVCEQKGLINASASTFFEA GVLVGAVAGVLSDFGRRELLLVAYVSTLV LGLASAASVSYVMFATTRILTGSALAGFTIIV MPLELEWILDVEHRTVAGVLSSTTFWGGVML LALWGYLERDWRWLLLAVTLPCAPGILS.UW VPESARWLLTQGHVKEAHRYLLICARLINGR PVCEDEFSGEAVSKVAAGERVVRRFSYLDLF RTPRLHHISLCCVVVWFGVNFSYYGLSDVS GLGINVYGTQLIFGAWEJRSKLLVYLSVRYA GRRLTQAGTLLGTALAFGTRLLVSSVRYA GRRLTQAGTLAGTALAFGTRLLVSSVRYA GRRLTQAGTLAGTALAFGTRLLVSSVRYA GRRLTQAGTLAGTALAFGTRLLVSSVRYA GRRLTQAGTLAGTALAFGTRLLVSSVRYA GRRLTQAGTLAGTALAFGTRLLVSSVRYA GRRLTQAGTLAGTALAFGTRLLVSSVRYA GRRLTQAGTLAGTALAFGTRLLVSSVRYA GRRLTQAGTLAGTALAFGTRLLVSSVRYA GRRLTQAGTLAGTALAFGTRLLVSSVRYA GRRLTQAGTLAGTALAFGTRLVSVRYA GRRLTQAGTLAGTALAFGTRLLVSSVRYA GRRLTQAGTLAGTALAFGTRLVSVRYA GRRLTQAGTLAGTALAFGTRLVSVRYA GRRLTQAGTLAGTALAFGTRLVSVRYA GRRLTQAGTLAGTALAFGTRLVSVRYA GRRLTQAGTLAGTALAFGTRLVSVRYA GRRLTQAGTLAGTALAFGTRLVSVRYA GRRLTQAGTLAGTALAFGTRLVSVRYA GRRLTVANSSAFTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGT	1030	2400	^	0730	3	1000	WVSSMGFEELLEQVGGFGPFQLRNVALLALP
NTILGERQSRGELEDEATTYPCSQWEYDM SERSSTIATESQWDIA/CGQGLINRAASTFFFA GYLVGAVAFGYLSDRFGRRRLLL/AYVSTLV LGLASAASYSYMPATITIGSALAGFTIIV MPLELEWIDVEHRTVAGVLSSTFWTGGVML LALVGYLERDWRWLLAVTL-CAPGILSLWW VPESARWLL/GGHYKAHRYLLHCARLINGR PVCEDSFSGEAYSKVAGERVYRRSYLDLY GLGINVYGTQLIFGAVELPSKLLVYLSVRYA GRRLTOAGTLIGTALAFGTLILVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALDGWUS LPALTYGGIALAAGTALLLPETRQAQLPETT QDVERKSAPTSLQEEEMMKQVQN TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALDGWUS LPALTYGGIALAAGTALLLPETRQAQLPETT QDVERKSAPTSLQEEEMMKQVQN GRRTTOAGTALLTPETRQAQLPETT QDVERKSAPTSLQEEEMMKQVQN GRRTTOAGTALLTPETRQAQLPETT QDVERKSAPTSLQEEEMMKQVGN GRRTTOAGTALLTPETRQAQLPETT QDVERKSAPTSLQEEEMMKQVGN GRRTTOAGTALLLAAGTALLLPETRQAQLPETT QDVERKSAPTSLQEEEMMKQVGN GRRTTOAGTALLTPETRQAQLPETT GODGVKVEKELENTEQPVGGGREVVEHEV TGNINSDPLIELCQCPLCDGSREGLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLISIEDGGSNCAVCGRALSSPGSLGG HLLISIEDGGSNCAVCGRALSSPGSLGG HLLISIEDGGSNCAVCGRALSSPGSLGG HLLISIEDGGSNCAVCGRALSSPGSLGG FENDAMSLKRANSTPEKRQARLQRDR QNEPLEVELQRLEBERTAKKSRRDNETPEERE VRRMRDREAKRLGRMQETDEGGRARLQRDR AGILLVONCAAYRKLLEAQTPSYRWALRR QNEPLEVELQRLEBERTAKKSRRDNETPEERE VRRMRDREAKRLGRMQETDEGGRARLQRDR AGILLVONCAAYRKLLEAQTPSYRWALRR QNEPLEVELQRLEBERTAKKSRRDNETPEERE VRRMRDREAKRLGRWGTDEGGRARLQRDR AGILLVONCAAYRKLLEAQTPSYRWALRR QNEPLEVELQRLEBERTAKKSRRDNETPEERE VRRMRDREAKRLGRARLAAFDNETP QLPYSOVELDSQLLGKMAFEGNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSSGYGA HGSKHARRAAPDPPFLEDDTSGGYSSOPGGY PATGADVARSVNHLLLGDWANAYASSI ASHGKDMVHKELLBETVSYSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSSKLLYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSSKLLYTHAUTSL ATVRSDLSTFHLLAYSGVYSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSKLLYTHAUT ATVRSDLSTFHLLAYSGVYSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSKLLYTHAUT ATVRSDLSTFHLLAYSGVYSKLKYFAUDTA YVAKKLGLLVFPYTHQNWEVQYSKLLTILGLYL ATVRSDLSTFHLLAYSGVYSKLKYFAUDTA YVAKKLGLLVFPYTHQNWEVQYSKLLTILGLYL ATVRSDLSTFHLLAYSGVYSKLKYFAUDTA YVAKKLGLLVFPYTHQNWEVQYSKLKTAAL GPDSMGGPYPRQRLQLYLTYCHW SNLKKAGQGGKKRKEKEEMBELRSY SSLMKVANSSNQDGNDSDEFM NEGTTLACRGGKGKEE						ì	HODAWA EALA DREDDOTT COOL DEALTHOUSE
SERSSTIATESQWDLVCEQKGLINASTFFFA GVLVGAVAGYGISDRFGRRILLIVAYVSTLV LGLASAASVSYVMPATRRITIGSALAGFTIIV MPLELEWIDVEHRTVAGVILAGTIIV MPLELEWIDVEHRTVAGVILAGTIIV MPLELEWIDVEHRTVAGVILAGTIIV MPLELEWIDVEHRTVAGVILAGTIIV MPLELEWIDVEHRTVAGVILAGTIIV MPLELEWIDVEHRTVAGVILAGTIITUSALAGFTIIV MPLELEWIDVEHRTVAGVILAGTIIGUSLAGFTIIV MPLELEWIDVEHRTVAGVILAGTIIGUSLAGFTIIV MPLELEWIDVEHRTVAGVILAGTIIGUSLAGVIIV RITRILRHISLCCVVVWFGVIRSYYJGLIDVS GIGLINVYGTQLIGAGVEPSKLLYVLSVRYA GRRITDAGTILGTALAFGTRILVSSDMKSWS TVLAVMGKAFSEAAFTTAYLTSELLYPTVIL QTGMGITALVGRIGGSLAFLAALLDGVWLS LPKLTYGGIALLAAGTALLIPETRQAQIPETI QDVERKSAFTSIQEEEMPMKQVQN LPKLTYGGIALLAAGTALLIPETRQAQIPETI TGYLNSDPILEICQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRASPISIAGIGG HLHISEDQRSNCAVCGARFTSIATFNSEKLP EVINMESILTVINIEGPSSAEGKDIAFSPVYY AGILVCNNCAYRKILBAQTPSVRKWALRR QNEPLEVRIQRILERERTAKKSRRIDETPEREE VRRNRDREAKRIQRMGFDEQGARRIQRDR EAMBLKRANSTPEKRQARIJREREAKRIKRR LEKMDMMIRAQFQQDPSAMAALAEMNIT QLPVSGVELDSQLLGKMAFEEQNSSSLH RICHMAMAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA				1			MITTI GEEDOGD CELEDEDAM DOGO COMPLEY
GVLVGAVAFGYLSDRFGRRRLLLVAYVSTLV LGLASAASVSYVMFATITGSALAGFTIV MPLELEWLDVEHRTVAGYLSSTFWTGGVML LALVGYLIRDWRWLLLAPICAPGILSLWW VPESARWILTQGHVKEAHRYLLHCARLNGR PVCEDSFSQEAVSKVAGERVYRRSYLDLF RTRRLRHISLCCVVVWFGVINFSYYGLSLDVS GLGINVYGTQLLFGAVELPSKLLVYLSVRYA GRRLTQAGTILGTALAFGTRILVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QUFBLAVORLGGSLAALDDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEMFMKQVQN QDVERKSAPTSLQEEMFMKQVQN RETTPVAVSSAPSGDSEGDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGREVSHEV TGNLNSDPLBLECQCPLCOLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGA HILHISEOGSNCAVGRARLQRDG EAMMLKRANSTEKKQRARLQRDG EAMMLKRANSTEKKQRARLQRDG EAMMLKRANSTEKKQRARLQRDG EAMMLKRANSTEKKQRARLQRDG EAMMLKRANSTEKKQRARLQRDG EAMMLKRANSTEKKQRARLQRDG EAMMLKRANSTEKKQRARLQRDG EAMMLKRANSTEKKQRARLQRDG EAMMLKRANSTEKKQRARLQRDG EAMMLKRANSTEKKQRARLQRDG EAMMLKRANSTEKKQRARLQRDG EAMMLKRANSTEKKQRARLQRDG EAMMLKRANSTEKKQRARLQRDG EAMMLKRANSTEKKQRANALYHSGYGA ASSLVFVPTRRSGPSGTSAVAMAYHSGYGA HGSKHRRAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDFMANVAMAYGSS ASHGKDMVHKELHERVSVKLKYFFAVDTA YVAKKLGLLVFFYTHQNMEVQYSGRAPLGLIGHV YQHKKLGLLVFFYTHQNMEVQYSGRAPLGLIGHV YVAKKLGLLVFFYTHQNMEVQYSGRAPLGLIGHV ATVRSDLSTFHLLAYSGVYYVGMLSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLTAAL GPDSMGGPVPRQRQLYLTLGAAAFQELITY ATVRSDLSTFHLLAYSGVYYVGMLSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLTAAL GPDSMGGPVPRQRQLYLTLGAAAFQELITY ATVRSDLSTFHLLAYSGVYYVGMLSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLTAAL GPDSMGGPVPRQRQLYLTLGAAAFQELITY ATVRSDLSTFHLLAYSGVYYVGMLSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLTAAL GPDSMGGPVPRQRQLYLTLGAAAFQELITY ATVRSDLSTFHLAYSGVYYVGMLSVLTCH LFGSDGYYVALAWTSSALMYFIVRSLTGLGHV SILKXTADMDVGQIGFHRQKDVKIVTVEKK VNELNRLEKTKVERFPDLAAKEGCRBEER NEKAQQGEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 327 RERTTLACRSGVYFSSCSLGACKFTYCSMP			l			!	SEESCHATESONDI ACEORGI VIDA VOLCEO
LGLASAASYSYVMATAITRITGSALAGFTIIV   MPLELEWILDVERITVAGVLSSTFWTGGVML   LALVGYLIRDWRWILLAVTLPCAPGILSL WW   VPESARWILITQGHVKEAHRYLLHCARLINGR   PVCEDSPSQEAVSKVAAGERVVRRPSYLDLF   RTPRIRHISLCCVVVWFGVNFSYTGLSLDVS   GLGINVYQTQLLFGAVELPSKILVYLSVRYA   GRRITQAGTILIGTALAFGTRILVSSDMKSWS   TVLAVMGKAFSRAAFTTAYLFTSELYPTUR   QTGMGLTALVGRIGGGIAPLAALIDGVWLS   LPKLTYGGIALLAAGTALLIPSTQAQLPETI   QTGMGLTALVGRIGGGIAPLAALIDGVWLS   LPKLTYGGIALLAAGTALLIPSTQAQLPETI   QTGMGLTALVGRIGGGIAPLAALIDGVWLS   LPKLTYGGIALLAAGTALLIPSTQDEVSSHIPS   EEDGGVVVEKELEPTTQDEVSSHIPS   EEDGGVVEKELEPTTQDEVSSHIPS   EEDGGVVVEKELEPTTQDEVSSHIPS   EEDGGVVVEKELEPTTQDEVSSHIPS   EEDGGVVVEKELEPTTQDEVSSHIPS   EEDGGVVVEKELEPTTQDEVSSHIPS   EEDGGVVVEKELEPTTQDEVSSHIPS   EEDGGVVVEKELEPTTQDEVSSHIPS   EEDGGVVVEKELEPTTQDEVSSHIPS   EEDGGVVVEKELEPTTQDEVSSHIPS   EEDGGVVVEKELEPTTQDEVSSHIPS   EEDGGVVEKELEPTTQDEVSSHIPS   EEDGGVVEKELEPTTQDEVSSHIPS   EEDGGVVEKELEPTTQDEVSSHIPS   EEDGGVVEKELEPTTQDEVSSHIPS   EEDGGVVEKELEPTTQDEVSSHIPS   EEDGGVVEKELEPTTQDEVSSHIPS   EEDGGVVEKELEPTTQDEVSSHIPS   EEDGGVVEKELEPTTQDEVSSHIPS   EEDGGVVEKELEPTTQDEVSSHIPS   EEDGGVVEKELEPTTQDEVSSHIPS   EEDGGVVEKELEPTTQDEVSSHIPS   EEDGGVVEKELEPTTQDEVSSHIPS   EEDGGVVETTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	1	1	ł			j	GVI VGAVAFGVI SDDEGDDDI I I VAVVISTI V
MPLELEWI,DVERTVAGVILSSTIFWTGGVML     LALVGYLRDWRWLLAVTLPCAPGILS, LWW     VPESARWILITQGHVKEAHRYLLHCARLNGR     PVCEDSFSQEAVSKVAAGERVVRRPSYTJDLF     RTPRILRIBLCCVVVWFGVNFSYYGLSLDVS     GLGLNYYGTQLLFGAVELPSKILLYYLSVRYA     GRRITQAGTILLGTALGTRILLYSDMKSWS     TVLAVMGKAFSEARFTTAYLFTSELYPTVIR     QTGMGLTALVGRLGGSLAPTKQAQLFETI     QDVERKSAFTSLQEEMPMKQVQN     LPKLTYGGIALLAGGTALLIPETRQAQLFETI     QDVERKSAFTSLQEEMPMKQVQN     LPKLTYGGIALLAGGTALLIPETRQAQLFETI     QDVERKSAFTSLQEEMPMKQVQN     ERIPPVANSAPSGDSEGDEETTQDEVSSHTS     EEDGGVVKVEKELENTEQPVGGAEVVEHEV     TGNLNSDPILELQCJLCQDCGSREQLIAHV     YQHTAAVVSAKSYMCPVCGRALSSPGSLGR     HILHSEDQRSNCAVCGARTSHATPNSEKLP     EVLNMESLFTVHNEGPSSAEGKDIAFSPGVT     AGILLVCNNCAAVKLLEAQTFSVRK WALRR     QNEPLEVRLQRLERERLAKKSRRDNETPERER     VRMRDREAKRLQRMQETDEQRARRLQRDR     EAMRLKRANETPEKRQARLIRERAKKLKRR     QNEPLEVRLQRLERERLAKKSRRDNETPERER     VRMRDREAKRLQRMQETDEQRARRLQRDR     EAMRLKRANETPEKRQARLIRERAKKLKRR     QLPVSGVELDSQLLGKMAFEEQNSSLH     GLEKNDMMLRAQFGQDPSAMAALAAEMNFT     QLPVSGVELDSQLLGKMAFEEQNSSLH     GLEKNDMMLRAQFGQDPSAMAALAAEMNFT     QLPVSGVELDSQLLGKMAFEEQNSSLH     GLEKNDMMLRAQFGQDPSAMAALAAEMNFT     QLPVSGVELDSQLLGKMAFEEQNSSLH     GLEKNDMMLRAQFGQDPSAMAALAAEMNFT     QLPVSGVELDSQLLGKMAFEEQNSSLH     GLEKNDMMLRAQFGQFSGSGTASVAAMAYHSGYGA     HGSGRGRGRGGRGRGGGVSSQPGGGV     PATGADVAFSVNELLGDPMANVAMAYGSSI     ASHGKDMYHKELHRFVSVSKLKYFFAVDTA     YVAKKLGLVPPTHGMYGGLGYSQPGGGA     ASLUFYTRISLGTHALAVSGYKYVOMILSVLTGL     LFGSDGYYVALAWTSSALMYFIVRSLRAAL     GPDSMGGPVPRQRLJCLASTALVWWEVVASLLLLGLYL     ATVRSDLSTHILAVSGYKYVOMILSVLTGL     LFGSDGYYVALAWTSSALMYFIVRSLRAAL     GPDSMGGPVPRQRLJCLASTALVKNSTUGGKMNNIVVYVTBW     SNILKKTADMDVGQIGFHRQKDVKNTTVEKK     VNELNRLEKTKVERFPDLAAEKCREMDELRSY     SSILMKVENNMSSNQDGNDSDEFM     REKKAQIQEMKRREKEMKKREMDELRSY     SSILMKVENNMSSNQDGNDSDEFM     TEATTLACRISCWVFSRSGLGACKFTVCSMP     1054   2404   A 8769 344 527     REATTLACRISCWVFSRSGLGACKFTVCSMP			l			l .	LGLASAASVSYVMFAITRTI TGSAI AGETTU
LALVGYLIRDWRWILLIAVTLPCAPGILSI.WW   YPESARWILTQGANVERNYLTLICAAPGILSI.WW    YPESARWILTQGANVERNYLTLICAAPGILISI.WY    YPESARWILTQGANVERNYLTYGLISDWS    GIGINVYQTQLIFGAQNIFSYYGISLDWS    GIGINVYQTQLIFGAQNIFSYYGISLDWS    STVLAVMGKAFSEAAFTTAYLFISELYYTVLR    QTGMGLTALVGRIGGSLAPLAALLDGUWLS    LPKLTYGGIALLAAGTRILIPETRQAQLPETI    QDVERKSAPTSLQEEMPMKQVQN    LPKLTYGGIALLAAGTRILIPETRQAQLPETI    QDVERKSAPTSLQEEMPMKQVQN    GIRTPVAVSSAPGDSEGDEEETTQDEVSSHTS    EDGGVVKVEKELENTLPDVGGNEVVERIEV    TGNLNSDPLLELQCCPLCQLDCGSREQLIAHV    YQHTAAVVSAKSYMCPVCGRALSSFGSLGR    HILHSEDQRSNCAVCGARTISHATFNSEKLP    EVINMESIPTYHNEGPVGGNEVVERIEV    TGNLNSDPLLELQCPLCQLDCGSREQLIAHV    YQHTAAVVSAKSYMCPVCGRALSSFGSLGR    HILHSEDQRSNCAVCGARTISHATFNSEKLP    EVINMESIPTYHNEGPVGGNEVVERIEV    TGNLNSDPLLELQCPLCQLDCGSREQLIAHV    YQHTAAVVSAKSYMCPVCGRALSSFGSLGR    HILHSEDQRSNCAVCGARTISHATFNSEKLP    EVINMESIPTYHNEGPVGGNEVVERIEV    TGNLNSDPLLERGETAKKSRRONETPEERS    VRRINGDREAKRIQROPTOGCARTISHATFNSEKLP    EVINMESIPTYHNEGPVGRAMALAAAMINF    QNEPLEVRLQRLERGETAKKSRRONETPEERS    VRRINGDREAKRIQROPTSORAMALAAAMINFPEERS    VRRINGDREAKRIQROPTSORAMALAAAMINFPEERS    VRRINGDREAKRIQROPTSORAMALAAAMINFPEERS    ASHGKDMYHKELHREVSVSKLKYFFAVDTA    YVAKLGLUVPPYTHROGGYSSQCYSQCYSQCYSQCYSQCYSQCYSQCYSQCYSQCYSQ	ŀ		1	<u> </u>			MPLELEWLDVEHRTVAGVI.SSTFWTGGVMI
VPESARWILITOGHVKEAHRYILHCARLNOR PVEDSPSQEAVSLAGER-VVRPSYIDLIS GLGLNYYOTQLIFGA-VELPSKILVYISVRYA GRRITQAGTILGTALAFGTRILIVSSIMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYFTVLR OTGMGITALVGRIGALAFGTRILIVSSIMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYFTVLR OTGMGITALVGRIGALAFGTRILIVSSIMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYFTVLR OTGMGITALVGRIGALAFGTRILIVSSIMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYFTVLR OTGMGITALVGRIGALAFGTALLIPETRQAQLPETI ODVERKSAPTSLQEEMPMKQVQN  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSIDESGIDEETTQDEVSSHTS EEDGGVVKVEKELENTTQPEVGGREVVEHEV TONLNSDPLIELCOCLOLDGGSREQLIAHV YQHTAAVVSSASYMCPVCGRALSSPGSLGR HILHISEDQRSNCAVCGARPTSHATFNSEKLP EVLNMESLFTVHREGFSSAEGKDIAFSPVYT AGILLVCNNCAAYRKLLEAQTTSVRKWALRR QNEPLEVVLQRLEERTAKKSRRDHSTPEERE VRRMRDREAKRLQRMQETDEQRARRILQRDR EAMRIKRANETPEKRQARLIERERAKKLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNET QLPVSGVELDSQLLGKMAFEEQNSSSLH RIGHGGRDRRGGGRARRICGRRGGGGYPORGAA ASLVFVPTRRSGPSGTASVAAMAYHISGYGA ASLVFVPTRRSGPSGTASVAAMAYHISGYGA HGSKHRARAAPPPEDFDDTSGGYSQPGOY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLVPFYTHQNWEVQYSRCHAPLFP RQDLNAPDLYIPTHAPHTYVLLAGMALGIQK RPSEVLGLASTALUVVVMEVLALLIGLVL LFGSDGYYVALAWTISSALMYTFYSLRTAAL GPDSMGGPYPRGULYTLTGAAAFQPLIIV WLTPHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYVSNELLIGENEDDIPKE VLMDCAHLVKANISQCKMINNVVVYTFW SNILKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNERTKVERPDLAFEKECNDEERR NEKKAQIQEMKKREKEEMMKKREMDELIPKE VLMDCAHLVKANISQQCKMINNVVVVYTFW SNILKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNHERTKVERPDLAFEKECNDEERR NEKKAQIQEMKKREKEEMMKKREMDELIPKE VLMDCAHLVKANISQQCKMINVVVVVYTFW SNILKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNHERTKVERPDLAFEKECNDEERR NEKKAQIQEMKKREKEEMMEKKREMDELIPKE VLMDCAHLVKANISNQDCHDDSDEFM	ľ	1	i				LALVGYLIRDWRWLLLAVTLPCAPGILSLWW
PVCEDSFSQEAVSKVAAGERVVRRPSYLDLF RTPRLRHSL.CVVVFGVNFSYYGJSLDVS GLGLNVYQTQLLFGAVELPSKLLVYLSVRYA GRRLTQAGTILLGTALAFGTRLLVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYFTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWULS LPKLTYGGIALLAACTALLPETRQAQLFETT QDVERKSAFTSLQEEMPMKQVQN  1051 2401 A 8759 515 1625 EERTPVAVSSAPSGDSEGDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPYGGREVVEHEV TGNLNSDPLLELCQCPLCQLDGGSRQULAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLHSEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVINGEGPSSAEGKDIAFSSPVYP AGILLVCNNCAAYRKILEAQTPSVRFWALRR QNEPLEVELQRLERERTAKKSRRDNSTPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANSTPEKRQARLIREREAKRLKRR LEKMDMMIRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEQNSSSLH LEKMDMMIRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEGNSSSH ASSALVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPLFDDTSGGYSSQFGGY PATGADVAFSVNHILGDPMANVAMAYGSSI ASHGKDMVHKELHRSVSKIKKYFAADTA AVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFTTYVLAGMALGIGK RTSEPVLGLASTALLWVVVMEVLALLIGLVL ATVRSDLSTFHILAYSGYKYVGMILSVLTIGI LFGSDGYYVALAWTYVRSSLAMYFTVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLITY WLTTHLLVR  1053 2403 A 8768 2 712 RPPRVWYELRELSAAAPRWSHRTAPGIMVF YFTSSSWISSAYTTYMGKDKYPNESLIKHGENEDDIFKE VLMDCAHLYKANSIQGCKMNNVNVYTFW SNILKKTADMDVGQIGFHRQRDVKIVTVEKK VNEILNBLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEPFM	1						VPESARWLLTQGHVKEAHRYLLHCARLNGR
RTPRIRHISLCCVVVWFGVNFSYYGLSLDVS GLGLNVYGTQLLFGAVELPSKLLVYLSVRYA GRRLTQAGTLLGTALAFGTRLLVSSDMKSWS TVLAVMGKAFSEAAFTTAYLTTSELYPTVLR QTGMGLTALVORILGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN  1051 2401 A 8759 515 1625 EERTPVAVSSAPSGDSEGDEEETTQDEVSSHTS EEDGGVVKVEKELENTEQPYGGNEVVEHEV TGNLNSDPLELECQFLQVLOGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HILLHISEDQRSNCAVCGARTSTHATTNSEKLP EVLNNBESLPTVINEGFSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRX WALRR QNEPLEVRLQRLERETAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEGRARRLQRDR EAMRLKRANSTPEKKQARLJRRREAKRLKRR LEKMDMMLRAQFOPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARRGLGRYPGRGAA ASLVYPYTRRSGPSGTASVAAMAYHSGYGA HGSKIRARAAPDPPPLFDDTSGGYSSOPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRPVSVSKLKYFAVDTA YVAKKLGLLVFPYTHONWEVQYSRDAPLPP RQDLNAPDLYPTNHAUTTVALAGAMLGIQK RPSPEVLGLCASTALVWVWMEVLALLLGLYL ATVRSDLSTPHLLAYSGYKYVGMLSVLTGL LFGSGGYYVALAWTSSALMYFIVTSLRTAAL GPDSMGGPVPRGRLQLYLTLGAAAFQPLITY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDK VENEDLIKHGW PEDIWFIYDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYTTEW SNLKKTADMDVGQIGFHRGKOVKIVTVEKK VNELNRLEKTKVERFPDLAAEEGCRBERR NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVERMSSNQDGNDSDEFM	1	]		]			PVCEDSFSQEAVSKVAAGERVVRRPSYLDLF
GLGLNYYQTQLIFGAVELPSKLLVYLSVRYA GRRLTQAGTILAFGTRLLVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRIGGSLAFLAALLDGWVLS LFKLTYGGIALLAAGTALLIPETRQAQLPETI ODVERKSAPTSLQEEMPMKQVQN ERTPVAVSSAPSGDSEGDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELQQPLQLDCGGREQLIAHV YQHITAAVVSAKYMCPVGGRASSPGSLG HLLHISEDQRSNCAVCGARPTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPYTY AGGILVCNNCAAYKKLLEAQTPSVKKWALRR QNEPLEVRLQRLERETAKKSRRDNISTPEREE VRRMDREAKRIQROGTDEGARRALQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMIRAQFGQDESAMAALAAEMNFF QLPVSOVELDSQLIGKMAFFEGNSSSLH  1052 2402 A 8763 1106 70 RHGHGGBDRRGGGRVARPGGLGRYPGRGAA ASLVFYPTRRSPSOTASVAAMAYHSGYGA HGSKHRARAPDPPPLFDDTSGGYSSQPGGY PATGADVARSVNHLLIGDPMANVAMAYGSSI ASHGKOMYHKEHFVSVSKLKKYFAVDTA YVAKKLGLVPPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAPFITYVLLAGMALGIQK RPSPEVLGLCASTALVWVVMEVLALLIGLYL ATVRSDLSTPHLLAYSGYKYYGMILSVLTGL LFGSDGYYVALAWVSAMEVLALLIGLYL ATVRSDLSTPHLJAYSGYKYYGMILSVLTGL LFGSDGYYVALAWSAMAYHSYSAMPTYNSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLITY WLTPHLVR 1053 2403 A 8768 2 712 RPPRFWYPPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAAHYVIRLHKGENIDFIKE VLMDCAHLVKANSIQGCKMNNVNVYTTEW SNLKKTADMDVGQIGFHROKDVKIVTVEKK VNEILNRLEKTKEPPDLAAKRECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENNSSNQDGINSDEFM NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENNSSNQDGINSDEFM							RTPRLRHISLCCVVVWFGVNFSYYGLSLDVS
1051 2401 A 8759 515 1625 ERRYPAVSASPOSEGDEEPTYQDEVSSHTS EDGGVVKVEKELENTEQPVGGREVVEHEV TONLNSDPILELCQCPI_COLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGE EVLINMESLETVHNEGPSSAGKDLAFSPVTY AGILLVCNNCAAVKKLLEAUTPSVRKWALRR QNEPLEVRLQRILGRAFSRAKKSRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQALIRERAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIRERAKKLRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHISGYGA HGSKHRARAAPDPPPLFDDTSGGYSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMYHKELHFVYSVSKLKVFFAVDTA YVAKKLGLLVPTHQNWEVYJSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVSLRAAL GPDSMGGPVPRQRLQLVLTLGAAAFQPLITY WLTFHLVR 1053 2403 A 8768 2 712 RPFRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTITYMGKKYENEDLIKHGW PEDLWFHVDKLSAYSHYYLLIKGKENEDIPKE VLMDCAHLVKANSIQGCKMNIVNVVYTPW SINLKCTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAKEGENDEERE NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENNSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACKNSCWYFSRCSLGACKFTVCSMP	ł						GLGLNVYQTQLLFGAVELPSKLLVYLSVRYA
1051 2401 A 8759 515 1625 EIRTPVAVSAPSGEDEEMPMCQVON  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEETTODEVSSHTS EEDGGVVKVELLENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLLCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLHISEDQRSNCAVCGARFISHATPINSEKLP EVLNMESLPTVHNEGEPSAEGKDIAFSPPVYP AGIILVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPERCQARLIREREAKRLKRR LEKMDMIRAQFGQDPSAMAALAAEMNIFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRFGGRVARPGELGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRRAAPDPPPLFDDTSGGYSSOPGGY PATGADVAFSVNHILLGEPMANVAMAYGSSI ASHGKDMVHKELHRPVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYPTMAFITYVLLAGMALGIQK RFSPEVIGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTPHLLAYSGYKYVOMILSVLTGL LFGSDGYYVALAWTSSALMYPTVRSLRTAAL GPDSMGGFVPRQRLQLYLTLGAAAFQPLITY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMKKKYENEDLIKHGW PEDLIWFHVKLSSAHVYLRLHKGENEDPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNILKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKKYERFPDLAARKECRDREER NEKKAQQEMKREKEEMKKREMDELRSY SSLMKVENMSSNQDGRNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWYFSRCSLGACKPTVCSMP							GRRLTQAGTLLGTALAFGTRLLVSSDMKSWS
LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVEKSAPTSLQEEMPMKQVQN  1051 2401 A 8759 515 1625 ERTPVAVSSAPSGDSEGDEESTTQDEVSSHTS EEDGGVKVEKELENTEQPVGGREVVEHEV TGRLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLHISEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVINGEGPSSAGEKDLAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRK WALRR QNEPLEVRLQRLEERSTAKKSRRNDETPEERE VRRMRDREAKLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSOLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFYTHONWEVQYSRDAPLPP RQDLNAPDLYPTMAPTTYVLLAGMALGIQK RPSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFTVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAARQPLITY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSVNSSAYTTYMKOKDK YENEDLIKHGW PEDIWFHVDKLSSAHTYILHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNIKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKKTVERFPDLAARKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM				,			TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR
DOTER SAPTS LOGGEMPMK QVQN   DOTER SAPTS	1						QTGMGLTALVGRLGGSLAPLAALLDGVWLS
1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEETTÖDEVSSHTS EEDGGVKVKYGGNEVVEHEV TGNLNSDPILEILOQCPI.CQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HILLHSEDQRSNCAVCGARPTSHATTNSEKLP EVLNNESI.PTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAVRKILEAQTPSVRKWALIRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMEDREAKRI.QRMQETDEQRARRLQRDR EAMRI.KRANETPERRQARI.IREREAKRI.KRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLIGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASILVFVPTTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVARSVNHILIGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGILVFPYTHQNWEVQYSRDAPLPP RQDI.NAPDLYIPTMAFITYVI.LAGMALGIQK RPSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTPHLLAYSGYKYVOMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSI.RTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLITY WLTTHLUVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHYVI.RLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVQQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATITACRNSCWVFSSCSLGACKFIYCSMP	ĺ						
EEDGGVVVEKERLENTEQPVGGNEVVEHEV TGNLNSDPILELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLOR HILHESDQRSNCAVCGARFTSHATFNSEKLP EVINMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKILEAQTPSVRKWALRR QNEPLEVELQRLERERTAKKSRRDNETTEERE VRRMRDREAKRI.QRMQETDEQRARRLQRDR EAMRLKRANETTPEKRQARLIREREAKRI.KRR LEKMDMMRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSI.H  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASSLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPLFDDTSGGYSSQPGGY PATGADVARSVNEHLIGDPMANVAMAYGSSI ASSHGKDMVHKELHRFVSVSKI.KYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIITMAFITYVLLAGMALGIQK RPSPEVLGLASTALVWVVMEVLALLIGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSALMYFIVRSLRTAAL GPDSMGGPPRGRQLQYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPFRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHYYLRLHKGENEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYYTPW SNIKKTADMDVGQIGFHRQKDVKIVTVEKK VNELNRLEETKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDDSDEFM  1054 2404 A 8769 344 527 REATILACRSCWVFSSCSLGACKFIYCSMP	1051	2401	A	8759	515	1625	CDAEKYPALIST OF THE CONTROL OF THE C
TGNLNSDPILELCQCPLCOLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLHISEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILL VCNNCAAYRKLLEAQTFSVRK WALRR QNEPLEVRLORLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEGRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH QLPVSGVELDSQLLGKMAFEEQNSSSLH GSKHRAAPDPPPLFDDTSGGGRVSOPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMYHKELHRFVSVSKLKYFFAVDTA YVAKLGLLVPPYTHQNWEVQYSRDAPLPP RQDLNAPDLYTPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIITY WLITFHLVR  1053 2403 A 8768 2 712 RPFRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFIRQKDVKVTVTEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATILACKNSCWYFSRCSLGACKPTVCSMP				0.55	313	1025	
YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HILIHSEDQRSNCAVCGARPTSHATFNSEKLP EVLNIMESLPTVHRIEGFSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKILEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEGRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMIRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH QLPVSGVELDSQLLGKMAFEGGRVARPGGGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVPYTHQNWEVQYSRDAPLPP RQDLNAPDLYPTMAPTTYVLLAGMALGIQK RPSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYVPAQLAVTSSALMYFTVSLATAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLITY WILTFHLVR  1053 2403 A 8768 2 712 RPFRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNILKTADMDVGGIGFHRQKDKVITVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATILACKNSCWYFSRCSLGACKPTVCSMP	1						
HILLIHSEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKILEAQTPSVRK WALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVFTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHILLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTPHILAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNIVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWYFSRCSLGACKFTVCSMP							YOHTAAVVSAKSYMCPVCGRAI SSPGSI GP
EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRK WALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMIRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVHBLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVPPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RPSPEVLGLCASTALVWVVMEVLAILLIGLYL ATVRSDLSTFHLLAYSGYKYVOMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHYYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPW SILIKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATITLACRNSCWYFSRCSLGACKPTVCSMP	1 .				i		
AGILL VCNNCAAYRKLLEAQTPSVRK WALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGIDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITTYLLAGMALGIQK RPSPEVLGLCASTALVWVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHYYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWYFSRCSLGACKPTVCSMP	l			ì			
QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETIDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFTTYVLLAGMALGIQK RPSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTPHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFFISSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATILACKNSCWVFSRCSLGACKFTVCSMP	ł	}			ļ	ļ	AGILLVCNNCAAYRKLLEAOTPSVRKWALRR
VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNIF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMYHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVWMEVLAILLIGLYL ATVRSDLSTPHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNILEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATILACRNSCWVFSRCSLGACKPTVCSMP					1		QNEPLEVRLQRLERERTAKKSRRDNETPEERE
EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFQQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAPHTYVLLAGMALGIQK RPSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTPHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLHY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDXYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNIVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNELNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWYFSRCSLGACKPTVCSMP							VRRMRDREAKRLQRMQETDEORARRLORDR
QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RPSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVOBMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHYYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATILACRNSCWVFSRCSLGACKPTVCSMP					ì	ł	EAMRLKRANETPEKRQARLIREREAKRLKRR
1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RPSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPFRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFFDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP							LEKMDMMLRAQFGQDPSAMAALAAEMNFF
ASLVFVPTRRSGPSGTASVAAMAYHSGYGA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RPSPEVLGLCASTALVWVWMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1052	2400		0763	1106		QLPVSGVELDSQLLGKMAFEEQNSSSLH
HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RPSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTPHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHYYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTFW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKERMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1032	2402	A.	6/03	1100	/0	
PATGADVAFSVNHILIGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGILVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RPSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTPHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHYYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTFW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKERDDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			- 1				ASLYTYPIKKKSGPSGIASVAAMAYHSGYGA
ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAPITYVLLAGMALGIQK RPSPEVLGLCASTALVWVVMEVLAILLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YYFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKERMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP				, [	ł		DATE ADVABBANTI LOOP ASSESSED OF
YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPIMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YYFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			ļ	ļ	. ]	j	
RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP				ŀ	. 1		
RPSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTPHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYPIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKEMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			ł	ŀ	İ		RODI NAPDI VIPTMARITVUI I ACMAI CICIV
ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPFRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		1	ł	ľ	l	İ	RPSPEVI.GLCASTALVWVVMEVI ATTICT VI
LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP				•	i		ATVRSDI.STFHLI.AYSGVKVVOMII QVI TOI
GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		J	J	]	[	ļ	LFGSDGYYVALAWTSSAI MYPIVRSI RTAAT
1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			· .		ĺ	1	GPDSMGGPVPRORLOLYI.TI.GAAAFOPI IIV
1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKERMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			]		ļ		
YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1053	2403	A	8768	2	712	
PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		ļ	J	1	}	J	YFTSSSVNSSAYTTYMGKDKYENEDLIKHOW
VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	'		ļ	f	ì		PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE
SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		•	1				VLMDCAHLVKANSIQGCKMNNVNVVYTPW
VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		ł		1	ł	I	SNLKKTADMDVGQIGFHRQKDVKIVTVEKK
NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM  1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		l		- 1		ļ	VNEILNRLEKTKVERFPDLAAEKECRDREER
1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		ļ	l		İ	İ	NEKKAQIQEMKKREKEEMKKKREMDELRSY
1 STORY I STOR	1054	2404	<u>, —                                   </u>	9760			SSLMKVENMSSNQDGNDSDEFM
SLSRQGSQTLCLRLAEYCMESVDSQRLLLS	1034	2404	A	g/07	344	527	REATTLACRNSCWVFSRCSLGACKPTVCSMP
							SLSKQGSQTLCLRLAEYCMESVDSQRLLLS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alaminc C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1055	2405	A	8770	430	1104	QQESPAAGAARMNCKEGTDSSCGCRGNDEK KMLKCVVVGDGAVGKTCLLMSYANDAFPEE YVPTVFDHYAVTVTVGGKQHLLGLYDTAGQ EDYNQLRPLSYPNTDVFLICFSVVNPASYHNV QEEWVPELKDCMPHVPYVLIGTQIDLRDDPK TLARLLYMKEKPLTYEHGVKLAKAIGAQCYL ECSALTQKGLKAVFDEAILTIFHPKKKKKKCS EGHSCCSII
1056	2406	Α	8773	261	332	NPRIQLSGNSCCAGSCRVWLSEQ
1057	2407	Α	8778	3	477	PAGIRHEQARGADRMGKCRGLRTARKLRSH RRDQKWHDKQYKKAHLGTALKANPFGGAS HAKGIVLEKVGVEAKQPNSAIRKCVRVQLIK NGKKITAFVPNDGCLNFIEENDEVLVAGFGR KGHAVGDIPGVRFKVVKVANVSLLALYKGK KERPRS
1058	2408	A	8808	171	881	PGLSQEPSGSMETVVIVAIGVLATIFLASFAAL VLVCRQRYCRPRDLLQRYDSKPIVDLIGAME TQSEPSELELDDVVITNPHIEAILENEDWIEDA SGLMSHCIAILKICHTLTEKLVAMTMGSGAK MKTSASVSDIIVVAKRISPRVDDVVKSMYPPL DPKLLDARTTALLLSVSHLVLVTRNACHLTG GLDWIDQSLSAAEEHLEVLREAALASEPDKG LPGPEGFLQEQSAI
1059	2409	A	8809	246		MRLQGAIFVLLPHLGPILVWLFTRDHMSGWC EGPRMLSWCPFYKVLLLVQTAIYSVVGYASY LVWKDLGGGLGWPLALPLGLYAVQLTISWT VLVLFFTVHNPGLALLHLLLLYGLVVSTALI WHPINKLAALLLLPYLAWLTVTSALTYHLWR DSLCPVHQPQPTEKSD
1060	2410	A	8810	304	381	PKLSVYPLQSHHCLSEPFQSLVCCLA
1061	2411	A	8820	1673	848	SCKTENLLEMWWFQQGLSFLPSALVIWTSAA FIFSYITAVTLHHIDPALPYISDTGTVAPEKCLF GAMLNIAAVLCIATTYVRYKQVHALSPEENVI IKLNKAGLVLGILSCLGLSIVANFQKTTLFAA HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS GNFGTDLEQKLHWNPEDKGYVLHMITTAAE WSMSPSFFGFFLTYIRDFQKISLRVEANLHGL TLYDTAPCPINNERTRLLSRDI
1062	2412	A	8824		763	GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNITSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIPYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV
1063	2413	A'	8826	147	627	CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR
1064	2414	A	8835	2982	1869	LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE HLQVITTLRTAAKEMEEKISNLKEHLASKEVE

NO. of   No. of   N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Deptide   Seq							D=Aspartic Acid. E=Glutamic Acid
Seq	nuci-						F=Phenylalanine G=Glycine H=Hietidine
1065   2415   A   8841   3   663   AATAASLSPRGCRLTFSSDVGPSRAPPSRAPING   AVAILABLE   AVAI	cotide	1	ĺ	USSN			
1065	seq-			1			M=Methionine N=Assaragine P=Proline
minio acid residue of peptide sequence      peptide   sequence   peptide   se	uence		Į.	914			O=Ghitamine R=Arginine S=Serine
Persidue of peptide sequence   Persidue of peptide sequence   Persidue nucleotide insertion   Persidue nucleotide nucleotide nucleotide nucleotide insertion   Persidue nucleotide		1	l			of peptide	T=Threonine, V=Valine, W=Tryptophan
Popesible uncleotide deletion, \ \popsible sequence		İ		1			Y=Tyrosine, X=Unknown, *=Stop codon.
			l		peptide		
VAKLEKQLLERKAMTDAMYPRSSYELCQ   SLSEWSVALASKLKSVEKEKYHSEVYQIES   EVSQVRREKENIQTILKSKEQEVYBLLQKPQ   QAQEELAMKRYSISSSKLEDDKOKKDEMS   KEVIKLKEALNISQUSYSTSSSKEQEQQLEA   LQQQVKQLQNQLAECKKQHQEVISYYRMHL   LYAVQQQMDEDVQKVLKQLTMCKNQSKK   KEVIKLKEALNISQUSYSTSSSKEQEQQLEA   LQQQVKQLQNQLAECKKQHQEVISYYRMHL   LYAVQQQMDEDVQKVLKQLTMCKQSKK   AAATAASLSPRGCRLRTPSSDVGPSRAPPPSA   APLPTGRAQMSPSGRLCLLTIVGLIPTRGGTL   KUTTSSSSAATMBIQVFTRAPADVTELQP   TSTIPTWADBITQPQTQTQQLEGTDGH-VT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   DPTHEKSTKAAHTPIDTITLSEKPSSTDVQT   RRINGERVYSKKMSLKSERGHIPHYQDLL   CKKGCGYYGMPWQGCSKCWREEYHKAR   DPTHEKSTKAAHTPICTRONGTON   DETAIL   CKKGCGYYGMPWGGCSKCWREEYHKAR   DPTHEKSTKAAHTPICTRONGTON   DETAIL   CKKGCGYYGMPWGGCSKCWREEYHKAR   DPTHEKSTKAAHTPICTRONGTON   DETAIL   CKKGCGYYGMPWGGCSKCWREEYHKAR   DPTHEKSTKAAHTPICTRONGTON   DETAIL   CKKGCGYYGMPWGGCSKCWREEYHKAR   DQUECKTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTKAAHTPICTRONGTON   DETAIL   DPTHEKSTK							
SLESEVSVILASKLKESVKEKERVHSEVVQIR		•			T		
BYSQVKREKENIQTILLKSKEGEVNBLLQKEMS   REVITALKEALNSLSQLSYSTSSSKLEDGEN KKINEMS   KEVITALKEALNSLSQLSYSTSSSKRQSQOLE   LQQQVKQLQNQLACKKQRQKSVSYNRMH. LYAVQGQMDEDVQKVLKQILTMCKNQSQK   K	1					ļ	SLESEVSVLASKLKESVKEKEKVHSEVVOIRS
		]					EVSOVKREKENIOTLLKSKEOEVNELLOKFO
1065   2415   A   8841   3   663	1 .	İ		[			
LQQQVKQLQNQLAECKKQHQEVISYYRMHL   LYAVQQQMDEDPUQKVLKQLIKCKNQSQK   K	İ	•	Ì			ł	KEVTKLKEALNSLSQLSYSTSSSKROSQOLEA
LYAVQQQMDEDVQKVLKQILTMCKNQSQK   K	İ		Ì				LQQQVKQLQNQLAECKKQHQEVISVYRMHL
1065   2415   A   8841   3   663	1	]		]			LYAVQGQMDEDVQKVLKQILTMCKNOSOK
APLPTGRAQMSPSGRLCLITVGLILPTRGQTI KDTTSSSSADATIMDIQVPTRAPDAYTELQP SPPTPTWPADETPQPQTQTQQLGGTDGPLVT DPETHKSTKAAHPTDDTTILSERPSSTDVQT DPQTLKPSGFHEDPFFYDEHTRKRGLLVA AVLFTTGIILTSGKCRQLSRLCRNHGR RRRRGRVSRKKMSLSERRGIHVDQSDLL CKKGCGYYGRPAWQGFCSKCWREEYHKAR QKQQEDWELAERLQREEEAFASQSSQCA QSLTSKFEEKKTNBKTRKVTNTVKKFFSASSI VGSKKGIOHAKAPSPSINRQTSIETDRVSKEFE FLKTFHKTGQETYKQTKLFLEGMHYKRDLSIE EQSSCAQDP/HINVAERMQTRGKYPPERVEKI MQGEKYHMTRLTKYYPCEPTTDDEKKDLSIE EQSSCAQDP/HINVAERMQTRGKYPPERVEKI MQGEKYHMTRLTKYYPCEPTTDDEKKDLSIE EQSSCAQDP/HINVAERMQTRGKYPPERVEKI MQGEKYHMTRLTKYYPCEPTTDDEKKDLSIE EQSSCAQDP/HINVAERMQTRGKYPPERVEKI MQGEKYHMTRLTKYYPCEPTTDDEKKDLSIE EQSSCAQDP/HINVAERMQTRGKYPPERVEKI MQGEKYHMTRLTKYYPCEPTTDDEKKDLSIE EQSSCAQDP/HINVAERMQTRGKYPPERVEKI MQGEKYHMTRLTKYYPCEPTTDDEKKDLSIE EXDAQSLAISQSDFPRYMGORPFLQSNI QVTTRFCNPSRLMTGEDGYYFTINLCCAVAFIE KLDAQSLNLSQSDFPRYMGORPFRQGERS WSPDACLGVKQMYKNIDLLSQLNERGERM NEAKKLEKDLDWTDGIABEVDFRKQPERS WSPDACLGVKQMYKNIDLLSQLNERGERM NEAKKLEKDLDWTDGIABEVDFRKQPERS WSPDACLGVKQMYKNIDLLSQLNERGERM NEAKKLEKDLDWTDGIABEVDFRKQPERS WSPDACLGVKQMYKNIDLLSQLNERGERM NEAKKLEKDLDWTDGIABEVDFRKQPERS WSPDACLGVKQMYKNIDLLSQLNERGERM NEAKKLEKDLDWTDGIABEVDFRKQPERS WSPDACLGVKQMYKNIDLLSQLNERGERM NEAKKLEKDLDWTDGIABEVDFRKQPERVY NEAKHLEENDUT PPTRGGTYQGFPQGP PPQGGYPQGPPYPQGGYPQGP PPQGGYPQGPPYPQGGYPQGP PPQGGYPQGPPYPQGGYPQGP PPQGGYPQGPPYPQGGYPQGP PPQGGYPQGPPYPQGGYPQGP PPQGGYPQGPPYPQGGYPQGP PPQGGYPQGPPYPQGGYPQGP PPQGGYPQGPPYPQGGYPQGP PPQGGYPQGPPYPQGGYPQGP PPQGGYPQGPPYPQGGYPQGP PTQSPPPNPYPQADFQVPPGGPPDSPGHGNYQ LVLTLQLSVLLSCCGDFRRKHPWINL VALSVLTASLSVMYGMMSPYTANDMANG UTTAVCFTVVPSNQTREVPLLYSCCGDFRRKHPWINL VALSVLTASLSVMYGMMSPYTANDMANG UTTAVCFTVVPSNQTREVPLLYDINI FLYTLTIGGARE*PSSSSLC*R.RWHGWMSPCP WIGGASCTSHLSCCQAGPREKDASLQPSCMY TADTSIWTRCGISMAPLVPPPPRGTGKATHPC HILLSTHCCMSPVQQPTGTGGSTRSGEGLSQ EVRYHVPPPPAPQGPCVPPGGGYPQGP PUGGSSGLEVRYLNIKKAVYDRPTASII LNGEKLKVPPVRSOTI*QGCSWP DDKMGNTFFRERDCRYLVNIKGKADMAYDM URHFFM*SINLIMEETYLNIKAVYDRPTASII LNGEKLKVPPVRSOTI*QGCMLACGGMMLAQALICRH  1072 2422 A 8870 33 658 MESVISSKYEDQGTIFTDYLEETYLNIKAVHPSTNT							
API-ITGRAQMSPSGILCLITIVGILIPTRGGTL   KDITSSSSADATIMDIQVPTRADYTELQP   TSPIPTWPADETPQPQTQQQLEGTDPUYT   TSPIPTWPADETPQPQTQQQLEGTDPUYT   TSPIPTWPADETPQPQTQQQLEGTDPUYT   TSPIPTWPADETPQPQTQQQLEGTDPUYT   TSPIPTWPADETPQPQTQQLEGTDPUYT   TSPIPTWPADETPQPQTQQLEGTDPUYT   TSPIPTWPADETPQPQTQQLEGTDPUYT   TSPIPTWPADETPQPQTQQLEGTDPUYT   TSPIPTWPADETPQPQTQLEGTDPUYT   TSPIPTWPADETPQPQTTGGAGGAGGAGGATTTGGAGGAGGAGGATTTGGAGGAGG	1065	2415	Α	8841	3	663	AAATAASLSPRGCRLRTPSSDVGPSRAPPPSA
KDTTSSSSADATIMDIQYPTRAPDAYTELQP   TSPITPIWADETPQPQTQLGTIDEPLYT		i					APLPTGRAQMSPSGRLCLLTTVGLILPTRGOTL
DPETIHKSTKAAHPTDDTTTLSEPPSSTDVOT	1						KDTTSSSSADATIMDIQVPTRAPDAVYTELOP
DPETIHKSTKAAHPTDDTTTLSEPPSSTDVOT	1		ľ				TSPTPTWPADETPQPQTQTQQLEGTDGPLVT
DPQTILKPSGHEDDPFTYDEHTJRRGLLVA	f						DPETHKSTKAAHPTDDTTTLSERPSPSTDVQT
1066					;		DPQTLKPSGFHEDDPFFYDEHTLRKRGLLVA
RRRRGRVVSRKKMSLKSERRGHIVDQSDLL	1000	1	<u> </u>	00.50	2005		AVLFITGIILTSGKCRQLSRLCRNHCR
CKKGGGYYGHPAUQGFCSKCWREEYHKAR   QKQIQEDWELARDREEAPASSOSSOGA   QSLTFSKFEEKKTNEKTRKVITVKKFFSASSR   VGSKKEIQEAKAPSPSINRQTSIETDRVSKEFIE   FLKTFHIKTGGEIVKYOTKLFLEGMHYVKRDLSIE   EQSECAQDPYHNVAERMQTRGKVPPERVEKI   MODIEKYMTRLJVYCPCETTDDEKKDLAI   QKRIRALRWVTPQMLCVPVNEDIPEVSDMVV   KAITOIIEMDSKRVFRDKLIACITKCSKHIPHAI   KITKNEPASADDPLPTLJVIVLKGHPPELQSNI   QYTIRFCNPSRLMTGEDGYYFTNLCAVAFIE   KLDAQSLNLSQEDEPRYMSGQTSPRRQEARS   WSPDACLGVKQMYKNLDLLSQLNERQERIM   NEAKCLEKOLIDWTDGIAREVQDIVEKYPLEI   KPPNQPLAADSENVEMDKLPPPLQVYAG   KPPNQPLAADSENVEMDKLPPPLQVYAG   KPPNQPLAADSENVEMDKLPPPLQVYYAG   KPPNQPLAADSENVEMDKLPPPLQVYYAG   KPPNQPLAADSENVEMDKLPPPLQVYYAG   KPPNQPLAADSENVEMDKLPPPLQVYAG   KPPNQPPPPLQVPQGYPQGPPQQGPPQQGPPQQPPQGGPPQQPPQQGPPQQPPQQ	1000	2416	A	8853	3806	2204	FVGEQEGGCEAGAGRGAQTYPGEAGERWFG
QKQIQEDWELARIACRIC QREEEAPASSQSSQGA   QSLTFSKPEKKKTRKYTTVKKFFSASSR   VGSKKEIQEAKAPPSINRQTSIETDRVSKEFIE   FLKTFHKTGQEIYKQTIKLFLEGMHYKRDLSIE   EQSECAQDPYHNVAREMQTRGKVPPERVEKI   MDQIEKYIMTRLJKYVPFCPETTDDEKKDLAI   QKRIRALRWVTPQMLCVPVNEDIPEVSDMVV   KATTDIIEMDSKRVPRQMKLACITKCSKHIPNAI   KITKNEPASADDFLPTLJYVIKGNPPRLJOSNI   QVTIRFCNPSRLMTGEDGYYFFINLCAVAFIE   KLDAQSLNLSQEDFDRYMSQGTSPRKQEAES   WSPDACLGVKQMYKNLDLISQLNERQERIM   MEAKKLENDLIDWTDGIAREVQDIVEKYPLEI   KPPNQPLAAIDSENVENDKLPPPLQPQVYAG   SMEREVGCGWLVPJPEAFWEAEVGGSLEARS   LRQAWATKQDPISKKK   1069   2419   A 8855   1530   1583   PCRPGMECNSMISVHCNL   1069   2419   A 8857   1530   1583   PCRPGMECNSMISVHCNL   1069   2420   A 8866   293   1675   PYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGYPYQGYPQGPYPQGYPQG	ļ .						KKRRGRVVSRKKMSLKSERRGIHVDQSDLL
QSLTFSKFEEKKHEKTRKYTVKKFFSASSR   VGSKKEIQEAK APPSINKOTISETDRVSKEFIE   FLKTFHKTGQEIYKQTKLFLEGMHYKRDLSIE   EQSECAQDFYHNVAERMQTRGKYPFERVEKI   MQQIEKYMTKLI, VYCFETTDDEKKDLAI   QKRIRALRWITQMLCUPVNEDIPEVSDMVV   KAITDIEMDSRK PYKOFLACTIK CSKHIPNAI   QKRIRALRWITQMLCUPVNEDIPEVSDMVV   KAITDIEMDSRK PYKOFLACTIK CSKHIPNAI   QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE   KLDAQSLNLSQEDFDRYMSGQTSPRKQEAES   WSPDACLGVKQMYKNLDLLSQLNRRQERIM   NEAKLEKDLDWTYPIPAFWGAEVGGSLEARS   LKDAWAYKNDPISKAL   KPPNQPLAAIDSENVENDKLPPFLQPQVYAG   KPPNQPLAAIDSENVENDKLPPFLQPQVYAG   KPPNQPLAAIDSENVENDKLPPFLQPQVYAG   KPPNQPLAAIDSENVENDKLPPFLQPQVYAG   KPPNQWATKQDPISKAL   KPPNQPLAAIDSENVENDKLPPFLQPQVYAG   KPPNQWATKQDPISKAL   KPPNQPLAAIDSENVENDKLPPFLQPQVYAG   KPPNQWATKQDPISKAL   KPPNQWATKQDPI							CKKGCGYYGNPAWQGFCSKCWREEYHKAR
VGSKKEIGGKARPSPSINGTSIETDRVSKEFFE   FLKTFHKTGGEIYKQTKLFLEGMFYKRDLSIE   FLKTFHKTGGEIYKQTKLFLEGMFYKRDLSIE   EQSECAQDPYHNVAERMQTRGKYPPERVEKI     MDQIEKYIMTRLYKYVFCPETTDDEKKDLAI     QKRIRALRWVTPQMLCVPVMEDIPEVSDMVV     KAITDIIEMDSKRVPRDKLACITKCSKHIPNAI     KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI     QVITRFCNPSRLMTGEDGYYFTNLCCAVAFIE     KLDAGSLNLSQEDFDRYMSGQTSPRKQEAES     WSPDACLGVK.QMYKNLDLLSQLNERQERIM     NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI     KPPRQPLAAIDSENVENDKLPPFLQPQVYAG     LRQAWATKQDPISKKK     LRQAWATKQDPISKKK     1068					,		QKQIQEDWELAERLQREEEEAFASSQSSQGA
FLKITHKTIQGETYKQTIKLFLEGMHYKRDLSIE							QSLTFSKFEEKKTNEKTRKVTTVKKFFSASSR
EQSECAQDY SINVAERM OT THE REPORT OF THE RE			1				VGSKKEIQEAKAPSPSINRQTSIETDRVSKEFIE
MDQIEKYIMTRLYKYYFCPETTDDEKKDLAI   QKRIRALRWTPQMLCVPVNEDIPEVSDMVV   KAITDIEMDSKRYPRDMLCHCYPVNEDIPEVSDMVV   KAITDIEMDSKRYPRDMLACITKCSKHIPNAI   KITKNEPASADDILJTLYIVLKGNPPRLOSNI   QYITRFCNPSILMTGEDGYYFTNLCCAVAFIE   KLDAQSLNLSQEDFDRYMSGQTSPRKQEAES   WSPDACLGVKQMYXKNLDLLSQLDREQERIM   NEAKKLEKDLDWTDGJAREVQDIVEKYPLEI   KPPNQPLAAIDSENVENDKLPPPLQPQVYAG   SNMREVGCWLYVIPJAFWEAEVQGSLEARS   LRQAWATKQPISKKK   LRQAWATKQPISKK   LR							FOSE OF ODER DELA PROCESSION OF THE PROCESSION O
QKRIRALRWYTPQMLCVPVNEDIPEVSDMYV							MOOTER AIMAN ARKWOLKOK ALAKA FILI
KAITDIEMDSKR/PRDKLACITKCSKHIFNAI   KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI   QYITRICNPSRLMTGEDGYYFTNLCCAVAFIE   KLDAQSLNI.SQEDFDR YMSGQTSPRKQEAES   WSPDACLGVKQMYKNLDLLSQLNERQERIM   NEAKKLEKDLDWTDGIAREVQDIVEKYPLEI   KPPNQPLAAIDSENVENDKLPPPLQPQVYAG   NEAKKLEKDLDWTDGIAREVQDIVEKYPLEI   KPPNQPLAAIDSENVENDKLPPPLQPQVYAG   SNMREVGCGWLVPVIPAFWEAEVGGSLEARS   LRQAWATKQDPISKKK   LRQAWATKQDPISKK   LRQAWATKQDPIPQGGYPQGPYPQGGYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGGYGGGGGGGG		,	. J				OK DID AT DATATEONAL CARRAGED TO EXCED AND
KITKNEPASADDFLPTLYIVILKGNPPRLQSNI QYTIRFCNPSRLMTGEDGYYFTNLCCAVAFIE KLDAQSLNLSQEDFDRYMSGQTSPRKQEAES WSPDACLGVKQMYKNLDLLSQLNERQERIM NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI KPPNQPLAAIDSENVENDKLPPPLQPQVYAG KPPNQPLAAIDSENVENDKLPPPLQPQVYAG SNMREVGCGWLVPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK LEQUAWATKQDPISKKK LEQUAWATKQDPSKKK LEQUAWATKQDST LEQUAWATKQDS LEQUAWATKQDST LEQUAWATKQDST LEQUAWATKQDS LEQUAMATKQDS LEQUAWATKQDS LEQUAWATKQDS LEQUAWATKQDS LEQUAMATKQ	[		' I				KATTOHEMDEKBADBUKI VCLIKCEKTIERIVI
QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE   KLDAQSLNLSQEDFDRYMSQGTSPRKQEAES   WSPDACLGVKQMYKNLDLLSQLNERQERIM   NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI   KPPNQPLAAIDSENVENDKLPPLQPQVYAG   SNMREVGGWLVPVPAFWEAEVGGSLEARS   LRQAWATKQDPISKKK   LRQAWATKQDPISKKY   LRQAWATKQDPISKKY   LRQAWATKQDPISKKY   LRQAWATKQDPISKY   LRQ	1.		ľ		·		KITKNEPASADDELPTI IVIVI KGNPPRI OGNI
				- 1			OYITRECOPSRIMTGEDGYYETNI CCAVAFIE
WSPDACLGVKQMYKNLDLLSQLNERQERIM   NEAKKLEKDLIDWTDGIAREVQDIVEKYPLE    NEAKKLEKDLIDWTDGIAREVQDIVEKYPLE    KPPQOPLAAIDSENVENDKLPPPLQPQVYAG   SNMREVGCGWLVPVIPAFWEAEVGGSLEARS   LRQAWATKQDPISKKK   LYLTIQISK	ļ	}	J				KLDAOSLNI SOEDEDRYMSGOTSPRKOFAES
NEAKKLEKDLIDWTDGJAREVQDIVEKYPLEI   KPPNQPLAAIDSENVENDKLPPPLQPQVYAG	<b>!</b>		l				WSPDACLGVKOMYKNI.DLLSOLNEROERIM
1067   2417   A   8855   1372   1513   SNMREVGGWLVFVIPAFWEAEVGGSLEARS     1068   2418   A   8856   1530   1583   PCRPGMECNSMISVHCNL     1069   2419   A   8857   1530   1583   PCRPGMECNSMISVHCNL     1070   2420   A   8866   293   1675   PYPQGGYPQGPYPQGYPQGPYPQGPYPQGPYPQGPYPQG	l i		İ		i		NEAKKLEKDLIDWTDGIAREVODIVEKYPLEI
1067   2417   A   8855   1372   1513   SNMREVGCGWLVPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK     1068   2418   A   8856   1530   1583   PCRPGMECNSMISVHCNL     1069   2419   A   8857   1530   1583   PCRPGMECNSMISVHCNL     1070   2420   A   8866   293   1675   PYPQGGYPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQ				1			KPPNQPLAAIDSENVENDKLPPPLQPQVYAG
LRQAWATKQDPISKKK	1067	2417	A	8855	1372	1513	SNMREVGCGWLVPVIPAFWEAEVGGSLEARS
1069   2419   A   8857   1530   1583   PCRPGMECNSMISVHCNL							
1070 2420 A 8857 1530 1583 PCRPGMECNSMISVHCNL 1070 2420 A 8866 293 1675 PYPQGGYPQGPYPQGGYPQGP YPQSFPPNPYQQPQVPPQGGYPQGPYPQGGYPQGP YPQSFPPNPYQQPQVPPQGPYPQGGYPQGP YPQSFPPNPYQQPQVPPQGPYPQGGYPQGP YPQSFPPNPYQQPQVPPQGPYPQGGYPQGP YPQSFPPNPYQQPQVPPQGPYPQGGYPQGP YPQSFPPNPYQQPQVPPQGPYPQGGYPQGP YPQSFPPNPYQQPQVPPQGPYPQGGYPQGP YPQSFPPNPYQQPQVPPQGYPQGPYPQGPYPQUPNU EEGPPSYYDNQDFPATNWDDKSIRQAFIKVF UVLTLQSVTLSTVSVFTISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG IITAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFFILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLIGNKQLSLSPEEYVPAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPIPGTGGSTRSGEGLSQ EVRYHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNILYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHILKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH				8856	1530	1583	PCRPGMECNSMISVHCNL
YPQSFPPNPYGQPQVFPGQDPDSPQHGNYQ EEGPPSYYDNQDFPATNWDDKSIRQAFIRKVF LVLTLQLSVTLSTVSVFITVAEVKGFVRENV WTYYVSYAVFISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG IITAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRYHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE  1071 2421 A 8868 2 358 ARGNILYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHFMI*SINLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH				8857		1583	
YPQSPFPPNPYGQPQVFPGQDPDSPQHGNYQ EEGIPSYYDNQDFPATNWDDKSIRQAFIRKVF LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMMASFYNTEAVIMAVG IITAVCFTVVIFSMQTRYDFTSCMGVILVSM VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCFLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRYHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE  1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLFEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH	1070	2420	A	8866	293	1675	PYPQGGYPQGPYPQEGYPQGPYPQGGYPQGP
EEGPPSYYDNQDFPATNWDDKSIRQAFIRKVF LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG IITAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE  1071 2421 A 8868 2 358 ARGNILYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLFEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		1		ľ		·	YPQSPFPPNPYGQPQVFPGQDPDSPOHGNYO
LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG IITAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVPAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRYHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE  1071 2421 A 8868 2 358 ARGNILYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		1	ļ	ļ	ļ	j	EEGPPSYYDNQDFPATNWDDKSIRQAFIRKVF
WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG IITAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEBYVPAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE  1071 2421 A 8868 2 358 ARGNILYHLPRLCTKLNLRWFSASTLYDVQH DDKMGSNIFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHILKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH			1				LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV
VALSVLTASLSYMVGMIASFYNTEAVIMAVG IITAVCFTVVIPSMQTRYDFTSCMGVLLVSM VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVPAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE  1071 2421 A 8868 2 358 ARGNILYHLPRLCRKLNILRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHILKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		1	- 1				WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL
ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFIFAIL.CIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEBYVPAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE  1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SINLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH	j l	1	- 1				VALSVLTASLSYMVGMIASFYNTEAVIMAVG
VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEBYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE  1071 2421 A 8868 2 358 ARGNILYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH	1	1	l l	- 1	•		ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM
VDTQLLLGNKQLSLSPEBYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE  1071 2421 A 8868 2 358 ARGNILYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SINKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		. [	- 1		ļ		VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA
WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HILLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE  1071 2421 A 8868 2 358 ARGNILYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH	ļ j	J	1	J	_ [		VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI
TADTSIWTRCGHSMAPLVLPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE  1071 2421 A 8868 2 358 ARGNILYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH				1	1	ſ	
HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE  1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNILRWFSASTLYDVQH DDKMGSNIFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH			Ì	İ	1	1	WHGSASCTSPLSCPQAQPREKDASLQPSCMY
1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH			J		1		TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC
1071 2421 A 8868 2 358 ARGNILYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SI\KLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH	j		j	J		1	HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ
1071 2421 A 8868 2 358 ARGNILYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH			1		1	ĺ	EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS
DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH	<u> </u>					<u> </u>	
DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SI\KLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP  1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH	107]	2421	A	8868	2	358	
1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		J	1		ı	ļ	DDKMGSNTFFKRNDCRYVMISCKADMAYDN
1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH			1		İ	ļ	VRHPFMI*SI\KLIMEETYLNIIKAVYDRPTASII
GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH	1000						
GGTGPSSDAGWGCMLRCGQMMLAQALICRH	10/2	2422	A	8870	33	658	MESVLSKYEDQITIFTDYLEEYPDTDELVWIL
			- 1		1	Į	GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI
LGRDWSWEKQKEQPKEYQRILQCFLDRKDC	Ì		- 1		ļ	Ţ	
· · · · · · · · · · · · · · · · · · ·							LGRDWSWEKQKEQPKEYQRILQCFLDRKDC

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	Ì	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
scq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
İ	i			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	i	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	1	peptide		/=possible nucleotide deletion, \=possible
	<u> </u>	<u> </u>	<u> </u>	sequence	<u></u>	nucleotide insertion
1		İ	,			CYSIHQMAQMGVGEGKSIGEWVLGPNTV\AQ
						GV*KNLA/LFDEW/NSLGLVYVSM/DNPSGSIA
	-	ļ				RFPKKLCRVLPL\SADTAGLTGP
1073	2423	Α	8879	146	412	DFSV*GDVDIEVTCPICLQLLTEPLSLNCGLRL
	Ì	1				*QVCITA*IKESVIISGG*SSSPVCHTTFQPANL
1074	10404	ļ. —	-			RTSRYLPT*SIKSLGPDEPQEG
1074	2424	Α	8884	67	435	HLQGRSIRTLQLTGENEKNCEVSERIRRSGPW
1	İ	ľ				KEISFGDYICHTFQGDCWADRSPLHEAAAHG
1			l			RLLALKTLIAQGVNVNLWTL/DRVSSLHEACL
1075	2425	A	8896	1294	040	*GPVACAKPYWKMVPRHGGTVTGPPLLMV
10/3	2425	A	0090	1294	248	RSGDRNGLTHQLGGLSQGSRNQSYRSRSRSR
		i				SRERPSAPRGIPFASASSSVYYGSYSRPYGSDK
1	İ	ŧ	i			PWPSLLDKEREESLRQKRLSERERIGELGAPE
	-					VWGLSPKNPEPDSDEHTPVEDEEPKKSTTSAS
			İ	i		TSEEEKKKKSSRSKERSKKRRKKKSSKRKHK KYSEDSDSDSDSETDSSDEDNKRRAKKAKKK
		ĺ				EKKKHRSKKYKKKRSKKSRKESSDSSSKES
		ĺ				QEEFLENPWKDRTKAEEPSDLIGPEAPKTLTS
ĺ		ļ				QDDKPLNYGHALLPGEGAAMAEYVKAGKRI
ł	1					PRRGEIGLTR*RNCHHLNAQVM**VVSRHRR
						MEAVRTAKREPESTVLMRREPLHPFNPRRET
	1					KERE
1076	2426	Α	8899	146	789	GRSTEAEKEPAFDERTGKGRRLPRAGEFHG*E
ł		}				*APGPGPRSFQVSRKMPEE\PPGARKHPFSGKS
						FYLDLPAGKNLQFLTGAIQOLGGVIEGFLSKE
j :						VSYTVSSRREVKAESSGKSHRGCPSPSPSEVR
						VETSAMVDPKGSHPRPSRKPVDSVPLSRGKE
	[			l		LLQKAIRNQK**CTVQQLSHCRLY\GEKTTAK
1077	2427	_	8901	750		RSQREHVQQQSQEHGKWPDLKGPR
10//	2421	A	8901	352	3	AKIGAYKYIQELWRKKQSDVMHFLLRVRCW
						QYPALHRAGTEWQLSALHRAPRSTQPDKAC
						RLGYKAKQGYIIYRICVRRGGWKCPVPKAVT
1078	2428	A	8905	536	781	\YGKPVHHGVN*LKFAQSLQSVAEEQ ACPAENREVPEMAAGQAPHAGPGAGPGQPA
			0,05	330	701	PALPFAATPGSRGQALCRGGRRRQHLHGPLH
						RP*QAAPALHAGCQLAPHPPT
1079	2429	A	8912	121	376 ·	NLIWKLCVTERRLVILDNYDLASE/YEANKYI
					-/-	CNRIIQFKPGQDKYFTLGLPTGSTPL*CYPKLI
			1	i		EYNKNGHLSFKYVKTFSMDEY
1080	2430	A	8920	381	1788	SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG
		.		ŀ		GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY
		]		İ		DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL
		}		l	1	YAPICMEYGRVTLPCRRLCQRAYSECSKLME
		i	ĺ	ľ	ï	MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA
		į	J	1		GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY
		l	ļ	ļ	1	SFLHVRDCSPPCPNMYFRREBLSFARYFIGLIS
		į.	- 1			IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV
		i	i	ŀ	•	WHMMVSLIFF\IGFLLEDRVACNA\SIPAQYKA
	1			į		STVTQGSHNKACTMLFMILYFFTMAGSVWW
		ł	ŀ	1	Į.	VILTITWFLAAVPKWGSEAIEKKALLFHASA
	ľ	l			l	WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD
		l	J	ļ	i	VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR
	ĺ	[	l	ĺ	ſ	VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL
1081	2421	<del>,  </del>	9000		400	VVIGCYFYEQAYRGIWETTWIQERC
1001	2431	A	8922	56	420	EERTKMSTGPDVKATVGDISSDGNLNVAQEE
		l			×	CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG
	1	- [			ľ	TLANFVF\CSVRHGLALILQLCNFSIYTQQMN
1082	2432	$\overline{\mathbf{A}}$	8923	355	1070	LSIAIPAMVNNTAPPSQPNASTERPST
	2752		رسون	333	1079	PFGTPSSTMAVVKNKCLMKGGKKGVKKKVV

NO: of peptide entitle outline   peptide entitle   peptide   peptide   peptide   peptide   peptide   peptide   pertite   per	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	LA-ing odd company (A. Ale I. G. G. A.
Decidic   Seq.   USSN   O9496   Oscarion   Overspondis   Pisobeusini, KLysine, L-Lucsine,   Pisobeusini, KLysine, L-					4		Amino acid sequence (A=Alanine C=Cysteine,
USSN   Docation   Do			1100				E-Phonylalonia C-Chaire H. Witt Fra
Sequence				1			F=rnenylalanine, G=Glycine, H=Histidine,
1914   1916   1916   1917   1916   1917			1			to lest emiss	F-Isoleucine, K=Lysine, L=Leucine,
minino acid residue of peptide   p	1 -	401100	Ì	i .		cold maidus	M-Meuricine, N-Asparagine, P-Proline,
Persidue of   Peptide	ucirco		ì	914			Q=Glutamine, K=Arginine, S=Serine,
		Ì	Į	1	•		1=1 hreonine, v=Valine, W=Tryptophan,
	ı		ì	1		sequence	Y=1yrosine, X=Unknown, *=Stop codon,
GPFSKKDQYDYKAPAMPNIRNTGK/TLVART   QFTGUSDLAGULFYSVALDLONDVAPK   PREMITED VODKNCL TNFYGMDL TCDKICSMV   REWSTMEAHPUVKTTDGYPFILE POYPTKK   HINQUILSTBY A*HQQSRQQKKMMEMT*BY   DYFIRKYKMLENPGFERMELROGGSS   ITWPQPHIPSCTAMSEETIQSKLAAKKKIP   DYFIRKYKMLENPGFERMELROGGSS   ITWPQPHIPSCTAMSEETIQSKLAAKKKIP   WGAVQGSRAMSDLLILDLITLLILIMLIGH   AGYSQQLAGVAYSAGSPPIRKYFIVEPYGET   WGAVQGSRAMSDLLILDLITLLILIMLIGH   AGYSQQLAGVAYSAGSPPIRKYFIVEPYGET   WGAVQGSRAMSDLLILDLITLILIMLIGH   AGYSQQLAGVAYSAGSPPIRKYFIVEPYGET   WGAVQGSRAMSDLLILDLITLILIMLIGH   AGYSQQLAGVAYSAGSPPIRKYFIVEPYGET   WGAVQGSRAMSDLLILDLITLILIMLIGH   AGYSQQLAGVAYSAGSPPIRKYFIVEPYGET   WGAVQGSRAMSDLLILDLITLILIMLIGH   AGYSQQLAGVAYSAGSPPIRKYFIVEPYGET   WGAVQGSRAMSDLLILDLITLILIMLIGH   AGYSQQLAGVAYSAGSPPIRKYFIVEPYGET   WGAVQGSRAMSDLLILDLITLILIMLIGH   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWQKKTINLRY   WGGRKVQFWKKWCKKTINLWFFILM   WGGRKVQFWKKWCKKTINLWFFILM   WGGRKVQFWKKWCKKTINLWFFILM   WGGRKVGFWKKWCKTINLTY   WGGRKVGFWKKWCKTINLTY   WGGRKVGFWKKWCKTINLTY   WGGRKVGFWKKWCKTINLTY   WGGRKVGFWKKWCKTINLTY   WGGRKVGFWKKWCKTINLTY   WGGRKVGFWKKWCKTINLTY   WGGRKVGFWKKWCKTINLTY   WGGRKVGFWKKWCKTINLTY   WGGRKVGFWKKWCKTINLTY   WGGRKVGFWKKWCKTINLTY   WGGRKVGFWKKWCKTINLTY   WGGRKVGFWKKWCKTINLTY   WGGRKYGFWKKWCKTINLTY   WGGRKYGFWKKWCKTINLTY   WGGRKYGFWKKWCKTINLTY   WGGRKYGFWKKWCKTINLTY   WGGRKYGFWKKWCKTINLTY   WGGRKYGFWKKWCKTINLTY   WGGRKYGFWKKWCKTINLTY   WGGRYGFWKKWCKTINLTY   WGGRYGFWKKWCKTINLTY   WGGRYGFWKKTINLTY   WGGRYGFWKKTINLTY   WGGRYGFWKTINLTY   WGGRYGFWKTINLTY   WGGRYGFWKTINLTY   WGGRYGFWKTINLTY   WGGRYGFWKTINLTY   WGGRYGFWKTINLTY   WGGRYGFWKTINLTY   WGGRYGFWKTINLTY   WGGRYGFWKTINLTY   WGGRYGFWKTINLTY   WGGRYGFWKTINLTY   WGGRYGFWKTINLTY   WGGRYGFWKTINLHOPSKYPKV   WGGRYGFWKTINLHOPSKYPK   WGGRYGFWKTINLHOPSKYPK   WGG	1	ł	ł	ł	,	1	/=possible nucleotide deletion, \=possible
QGTQIASDGLKGLI_FEVSLADLQNDEVAFER   FELTITEDVQDKNCLTNFYGMDLTCDKICSMW   EKWSTMEAHVDWKTTIDGVICSMW   EKWSTMEAHVDWKTTIDGVICSMW   EKWSTMEAHVDWKTTIDGVICSMW   EKWSTMEAHVDWKTTIDGVICSMW   EKWSTMEAHVDWKTTIDGVICSMW   EKWSTMEAHVDWKTTIDGVICSMW   EKWSTMEAHVDWKTTIDGVICHTEV   PWGAVQSRAQSRQJKDKDTEVCVCPYPLIDDVFRKYKMLEIPPGFRWELEJGGGSSS   DVFWPQFHIPSCPAMSETILOSKLAAAKKKIP   DVFRKYKMLEIPPGFRWELEJGGGSSS   WGAVQGSRAMSDLILLIDLITLIMLIGH   AGYSQQLAGVAVSAGSPPIRYKFPVEPYGET   GWLLTPSCSISFKCLSIAVHYDNAWF   EWGAVGSRAMSDLILLIDLITLIMLIGH   GWLTTPSCSISFKCLSIAVHYDNAWF   EWGAVGSRAMSDLILLIDLITLIMLIGH   GWLTTPSCSISFKCLSIAVHYDNAWF   EWGAVGSRAMSDLILLIDLITLIMLIGH   GWLTTPSCSISFKCLSIAVHYDNAWF   EWGAVGSRAMSDLILLIDLITLIMLIGH   GWLTTPSCSISFKVWQKKRTLARUP   EWGAVGSRAMSDLILLIDLITLIMLIGH   GWLTTPSCSISFKVWQKKRTLARUP   EWGAVGSRAMSDLILLIDLITLIMLIGH   GWLTTPSCSISFKVWQKKRTLARUP   EWGAVGSRAMSDRILLITLIMLIARUP   EWGAVGSPTGCWWGCKRLISFWKTT GSFAX   HTTPSTVDTAPISGNYPKRMSKKGQFTCARK   "HTTPSTVDTAPISGNYPKRMSKKGQFTCARK "HTTPSTVDTAPISGNYPKRMSKKGQFTCARK "HTTPSTVDTAPISGNYPKRMSKKGQFTCARK "HTTPSTVDTAPISGNYPKRMSKGCGFTCARK "HTTPSTVDTAPISGNYPKRMSKGCGFTCARK "HTTPSTVDTAPISGNYPKRMSKGCGFTCARK "HTTPSTVDTAPISGNYPKRMSKGCGFTCARK "HTTPSTVDTAPISGNYPKRMSKGCGFTCARK "HTTPSTVDTAPISGNYPKRMSKGCGFTCARK "HTTPSTVDTAPISGNYPKRMSKGCGFTCARK "HTTPSTVDTAPISGNYPKRMSKGCGFTCARK "HTTPSTVDTAPISGNYPKRMSKGCGFTCARK "HTTPSTVDTAPISGNYPKRMSKGCGFTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGTCARK "HTTPSTVDTAPISGNYPKRMSKGCGGGTCARK "HTTPSTVDTAPISGNAM "HTTPSTVDTAPISGNAM "HTTPSTVDTAPISGNAM "HTTPSTVDTAP	<del></del>		<del> </del> -	<del> </del>	sequence		
FRLITEDVQDKNCLTNFYGMDLTCDKICKNK	1	Ì	j			1	GPFSKKDQYDVKAPAMFNIRNTGK/TLVART
BEKWSTMIEAHUDVKTTIDGYFFHLECVGFTKHENNOLLTSYS-MCQSRQIGKKMEIMTEV	Ì		i	Ì		İ	QGTQIASDGLKGLLFEVSLADLQNDEVAFRK
HNNQLKTSYA*HQGSRG(GKKMEIMTER)   1083	1	ĺ	ĺ	1	i		FKLITEDVQDKNCLTNFYGMDLTCDKICSMV
		ĺ	}	ł			
1083		l	1		ĺ		HNNQILKTSYA*HQQS/RQIQKKMMEIMT*EV
1083	1	1	l		ł	l	
1084   2434   A   8950   156   318	1000						DVFIRKVKMLENPGFER\MELRGGGSSS
1084	1083	2433	ΙΑ.	8948	28	385	LTWPQPHIPSCPAMSEETLQSKLAAAKKKLP
1084   2434   A	1		i '			ĺ	WGAVQGSRAMSDLLLLLLDLTLLLLLMLLGF
1084							AGYSGQLAGVAVSAGSPPI/RYKFHVEPYGET
1085			<u> </u>				GWLLT/ESCSISPKLCSIAVH*DNPAWF
1085	1084	2434	Α	8950	156	318	HYTPINTDTIENSENNKCW*GY*E\VGLIHHW
1086							WGGKRVQPFWKRVWQKRTLNLRV
MFILAPFTATIKGKQLTCPLVEREIDYNMYYS	1085	2435	Α	8956	16	413	HMGQLGYFIQCWWECKRLISF\WKTI*OSPAK
MFILAPFTATIKGKQLTCPLVEREIDYNMYYS	i I		!	}			*TIYTSYDTAIPIS/GI/YPKRMSSKCHOETCAR
HKYYKVRRIL*VTITHYIWVILNIIMFEIILW YSHKYY   YCLQVYAMCYYYICH   YERSVCAFHIYCIQITYYCLQVYAMCYYYICH   YYSYCYGGLICTCVCMDYICVCYQEFL   YERSVCAFHIYCIQITYYCLQVYAMCYYYICH   YYSYCYCMDYYICVCYQEFL   YYSKYCGGLICTCVCMDYICVCYQEFL   YYSKYCGGLICTCVCMDYICVCYQEFL   KKFLDMSNAN*STKKHDKLDLIKKTICSA   KYTYKRIKHPTDLEKMRNHLSDKD*YSGVY   YKDLSKLNRRKTE;S*YVKKWVKDLSRYFIKE   YISMENKHKIIFSTS   YISMENKHKIIFSTS   YISMENKHKIIFSTS   YSSWSTYPISPSSLEDLEATGTIGTLLSDMGVV   GVEDNAYTLEVNSRYMRAVGIM*HIL   NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV   GVEDNAYTLEVNSRYMRAVGIM*HIL   GGYPGGTQSVFLTGVLVSSVYYMMKMLHTR   LLIAALFIUQYWK QSKDHYI   LUIAALFIUQYWK QSKDHYI   LVSSRYKISKVIVVGDLSVGKTCLIRR*GGAG   AELGRVOPSLARWAGSRQHLVPSQVCKDS   FDKNYKAPIGADPEMERFEVLGIFF   GGYPGGTQSVFLTGVLVSSVYYMMKMLHTR   LVSSRYKISKVIVVGDLSVGKTCLIRR*GGAG   AELGRVOPSLARWAGSRQHLVPSQVCKDS   FDKNYKAPIGADPEMERFEVLGIFF   GGYPGGTGSVFLTGVHPHPHPRFRQQ   FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH   RAAHHHQHQQGFLGHGLVARVG   TCPHICTVVNYKELAEHIRSKYPGCTPTITVVD   AMCLRYWYTPSSUKGQWRBYFSALRDF   VKTFTAAGIKLIFFFDGMVEQDKRDEWYCRR   LKNNREISRIFHYNKSHKEQPGRNMFTIPSGLA   VFIRFALKTLGQETLGSLQEADYEVASYGLQ   HNCLGILGEDTDYLIYDTCYPFSISELCLESILD   TVMLCREKLCESILG LOSLD   TVMLCREKLCESIL	]		l		. •		MFILAPFTATIKGKQLTCPLVEERIDY\MWYS
1086	!		l				HKYYIKVKRNL*VTITH\TWVNLNILMFEIILW
1087   2437   A   8985   58   330			l				YSHKYY
1087	1086	2436	Α	8962	868	1026	H*KILQVGRAQRAHXSRL*SQLLRRLRHESHL
1087						× .	NPGARGCSEARLHRCTPAWTT
** **ERSVCAFHVČIQTYVCLQVJACMCVYYTČICM FVYSVYGCGLCTCVCMDVYICVCQEFL	1087	2437	Α	8985	58	330	
1088							*ERSVCAFHVCIOTYVCLOVYACMCVYYICM
1088	Ĺ						FVYSVYGCGLCTCVCMDVYICVCVOEFL.
1089	1088	2438	Α	8989	394	404	N*KWILHVNVRIOSIFF/IKRNOK/INSHELKI.D
1089							KKFLDMMSNA*STKKHDKLD/LIKFKT/LCSA
1089	1						KYTVKRIKIHPTDLEKMLRNHLSDKD*YS/GV
1089 2439 A 8991 60 329 MALTPESPSSFPGLAATGSSVPEPPGGPNATL NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV GVEDNAYTLEVNSRYMRAVGIM*IHL 1090 2440 A 8996 2 351 SNITITLT*MKKYDNTFCW*GCGQIG/T/LYC WQESKPIQAPWSKIQQYLA*ISIHILFDPAFLFL GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI 1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AFLGRVGPSLARWAGSRSQHLVPSQVCKDS FDKNYKAPIGADFEMERFEVLGIFF 1092 2442 A 8999 548 811 SSFIKRHILFEDDWHQTTCCHHPHHPRP*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFDGMVCDKRDEWVKRR LKNNREISRIFHYKSHKEQPGRNMFIPFSGLA VFTRFALKTLQQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYPSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGRDII PEGMFESFRYKCLSSYTSVKENIPDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTIK**YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPMCTOSEPRQEVPMCTOSEP RQEVPMYTOSEPRQEVPMCTDSEP	}						YKDLSKLNRRKTE/S*/VKKWVKDI.SRYFIKE
1090 2440 A 8996 2 351 SNITITIT*MKKYDNTFCW*GGGGGGT/T/IYC WQESKFIQAFWSKIQQYLA*ISIHILFDPAFIFI GGYPGGTQSVFITISYSSYFYNMKMLHTR LLIALAFIIVQYWKQSKDHYI 1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AFLGRVGPSLARWAGSRSQHLVPSQVVCKDS FDKNYKAPIGADFEMERFEVLGIFF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPF*RCQ FHIFTYSVQNSISFSLSVSSSIPDRPDHEVHQH RAAHHIQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALIGLQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWTPESWICGGQWREYFSALRDF VKTTTAAGIKLIFFDDWVQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFTIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGI.CVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRTNLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQALNPGGKFPCVQALNPGGKFPCVQALNPGGFKFPCVQALPPGFFFRQEVPMCTGFFFRAGFFFAUNTGFFFRAGFFFFAUNTGFFFRAGFFFAUNTGFFFRAGFFFAUNTGFFFAUNTGFFFAUNTGFFFAUNTGFFFAUNTGFFFAUNTGFFFAUNTG							
NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV GVEDNAYTLEVNSRYMRAVGIM*HIL  1090 2440 A 8996 2 351 SITTITL**MKKYDNTFCW*GCGQIG/T/IYC WQESKFIQAFWSKIQQYLA*ISIHILFDPAFLFL GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI  1091 2441 A 8997 97 456 YPLFVCSYLSGPRGEHWNSLGGKSSCPLFLFT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQVCKDS FDKNYKAPIGADFEMERFEVLGIFF  1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHP\F*RCQ FHLFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG ALGLQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYIPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLQGETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLJYDTCPYFSISELCLESLD TVMLCEKLCESLGLCVADLPHLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEELPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTTPNGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEPRQEVPMCTDSEP	1089	2439	A	8991	60	329	
GVEDNAYTLEVNSRYMRAVGIM*IHL							NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV
1090   2440   A   8996   2   351   SNITITLT*MKKYDNTFCW*GCQIG/T/LIYC WQESKFIQAFWSKIQQYLA*ISHILIPDPAFLFL GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI   LLIAALFIIVQYWKQSKDHYI   LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLAR WAGSRSQHLVPSQVCKDS FDKNYKAPIGADFEMERFEVLGIPF   SFIKRHILIFEDDWHQTTCCHHPHHPF*RCQ FHIFYVSVQNISIPSLSVSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG   RAAHHQHGQGPLGHGLVARVG   AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFTDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGI.CVADLPLLACLLGNDII   PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEELPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTFNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPMCTDSEPRQEVPMCTDSEP RGEVPMYTGSEPRQEVPMY			1				
WQESKFIQAFWSKIQQYLA*ISIHILFDPAFLFL GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLLAALFIIVQYWKQSKDHYI  1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*OGAG ABLGRVGPSLARWAGSRSQHLVFSQVVCKDS FDKNYKAPIGADFEMERFEVLGIPF  1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPLF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHIHIQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/NLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKPPCVQALNP GEKFPCHI/PEPRQEVPTCTGPEPRQEVP MCTGPEARQEVPMCTGSEPRQEVPMCTGSEPR RQEVPMYTGSEPRQEVPMCTGSEPR	1090	2440	Α	8996	2	351	SNITTILT*MKKYDNTFCW*GCGOIG/T/LIYC
GGYPGCTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI  1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS FDKNYKAPIGADFEMERFEVLGIPF  1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPFFRCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHIHIQHGQCPLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGICVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKPPCVQMLNPGKKPPCVQALNP GEKFPCHIPPEPRQEVPMCTGSEPRQEVPMCTGSEP PESRREVPMCSDPEPRQEVPMCTGSEPRQEVPMY MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP			. •				WOESKFIOAFWSKIOOYLA*ISIHILFDPAFI.FI.
LLIAALFIIVQYWKQSKDHYI					ļ		GGYPGGTOSVFLTGVLVSSVFYNMKMLHTR
1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISK VIVVGDLSVGKTCLINR*GGAG AELGRQFPSLARWAGSRSQHLVPSQVVCKDS FDKNYKAPIGADFEMERFEVLGIPF 1092 2442 A 8999. 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPVP*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL*RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTTPNPGKKPPCVQMLNPGKKFPCVQALNP GEKFPCHIJ/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPERRQEVPMCTGPEPRQEVPMCTGPEPRQEVPMCTGPEPRQEVPMCTGPEPRQEVPMCTGPEPRQEVPMCTGPEPRQEVPMCTGSEPRQEVPMCTGPERRQEVPMCTGSEPRQEVPMCTGPERRQ					ł		
LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQVCKDS FDKNYKAPIGADFEMERFEVLGIPF  SSFIKRHILIFEDDWHQTTCCHHPHHPPF*RCQ FHIFYVSVQNSISFSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALGIQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGINDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPTCSDPEPRQEVPTCTG PESREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP	1091	2441	A	8997	97	456	YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT
AELGRVGPSLARWAGSRSQHLVPSQIVCKDS FDKNYKAPIGADFEMERFEVLGIPF  1092 2442 A 8999. 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPVF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHIQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPILACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMYT				1			
FDKNYKAPIGADFEMERFEVLGIPF			1				
1092 2442 A 8999. 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPVF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEELPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTGDEPRQEVP MCTGPEBRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTGPEPRQEVP RQEVPMYTGSEPRQEVPMCTGPEPRQEVP RQEVPMYTGSEPRQEVPMCTGPERQEVPMY		- 1	ł	1			
FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEELPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKPPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTGSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP	1092	2442	A	8999.	548	811	
RAAHHHQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEELPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTGSEPRQEVPMYT				]			FHIFYVSVONSISPSI SVSSSHPDRPDHEVHOH
ALIGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGPSKNPKV**LWTNK*YP RVQTPNPGKKPPCVQMLNPGKKPPCVQALNP GEKFPCIHIPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY		j	J				
TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGPSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHL/PEPRQEVPTCTGG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMYTGPESRQEVPMY	1093	2443	A	9002	3	2745	
AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGPSKNPKV**LWTNK*YP RVQTPNPGKKPPCVQMLNPGKKPPCVQALNP GEKFPCIHIPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	1						TCPHICTVVNEKEI VERIEBGKASCHIOAKOPOOLA
VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGPSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHL/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTGSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	i f	1		ļ	×-		
LKNNREISRIFHYIKSHKEQFGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY				Ì			
VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHL/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTGSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY		i				1	I KUNDETEDIETATIVETIKEODODAD GEGOOTA
HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHL/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTGSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	1		[	ſ	Ĭ	í	VETDEAL VIII COPTI COL OPADATIVA OTICI O
TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKPPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTGSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	. 1		Ī	ŀ			THICLGII GEDTOVI TVDYCDVEGGET CT POT D
PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTGSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY		ļ		ł	i		TVMI CREKI CESI CI CIVADI RILLACI COMP
VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	ł	ı	l	ł	1	ł	DECAME CED VY CI COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED VY COVTOUVEN DECAME CED V
*RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY						į	VEDUICEVII VI VOOREVE PRII DI ARTICOSSE
RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY			ł			i	*DAGREETETAR LATERATION
GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	į		i				PAGE TRANSPORTED TO A STREET OF THE PROPERTY O
PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	.	ļ	ŀ	1	l		RVQ1PNPUKAPPCVQMLNPGKKFPCVQALNP
MCTGPEARQEVPMCTDSEP RQEVPMYTGFESRQEVPMY RQEVPMYTGSEPRQEVPMYTGFESRQEVPMY	ľ	i	ď		Ì	- 1	GEKTYCIHUYEPKQEVYICSDPEPRQEVPTCTG
RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY		ì			1	į	PESKKEVPMCSDPEPRQEVP
		I				ļ	MCIGPEARQEVPMCTDSEP
TGPESRQEVLIRTDPESRQEIMCTGHESKQEV		ļ	i		J	j	
				<u>-</u>			IGPESKQEVLIK IDPESKQEIMCTGHESKQEV

SEO ID	SEQ ID	Met	Loro	Predicted	1 1 1 1 1 1 1 1	
NO: of	NO: of	hod	SEQ		Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence	1	09/496		corresponding	I=Isoleucine, K=Lysine, L=Leucine,
uence	) delice	)	914	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
ucucc			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ŀ				residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ľ	i	ł		,	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l		peptide		/=possible nucleotide deletion, \-possible
<b></b>	<del> </del>	<del> </del>	<del>  </del>	sequence		nucleotide insertion
1	ł		Ì			PICTOPISKQEDSMCTHAEINQKLPVATDFEFK
1	ļ	l			i	LEALMCTNPEIKQEDPTNVGPEVKQQVTMVS
İ	1					DTEILKVARTHHVQAESYLVYNIMSSGEIECS
1		ł		1	1	NTLEDELDQALPSQAFIYRPIRQRVYSLLLED
1 .	ŀ	l	1	1	ļ	CQDVTSTCLAVKEWFVYPGNPLRHPDLVRPL
	i		ļ		1	QMTIPGGTPSLKILWLNQEPEIQVRRLDTLLA
				ŀ		CFNLSSSREELQAVESPFQALCCLLIYLFVQV
i		1	ł	}		DTLCLEDLHAFIAQALCLQGKSTSQLVNLQP
		İ				DYINPRAVQLGSLLVRGLTTLVLVNSACGFP
	1	I				WKTSDFMPWNVFDGKLFHQKYLQSEKGYA
1				í		VEVL/CRTK*ISAHQIPQPEGSRLQGLHEGEQT
1004	244	L				HHWPSPLGLTPRREVGKTGLQLPQDGLWV
1094	2444	Α	9021	97	834	AREACRAKTDFPGRRFRLWPSCCCRVIVGAE
				'		T*H\MAEPVSPLKHFVLAKKAITAIFDQLLEFV
	1					TEGSHFVEATYKNPELDRIATEDDLVEMQGY
1						KDKLSIIGEVLSRRHMKVAFFGRTSSGKSSVI
	j					NAMLWDKVLPSGIGHTINCFLSVEGTDGDKA
	!					YLMTEGSDEKKSVKTVNQLAHALHMDKDLK
						AGCLVRVFWPKAKCALLRDDLVLVDGPGTD
1000						VTTELDSWIDKFCTKSSTREITNSGSDT
1095	2445	Α	9022	1	537	LVLNSRVEDFVPPEGAGRTLPFALRPLAACW
J						LLHRRARRSSALCPRPRSWGVSGGEGAGARE
]						P*ITSSSCCLSAA/SHLSIQSPNMAGARRRIRPQ
1						LAKEKIEGCHICTSVTPGEPQVFLGKDKAFTF
-						DYVFDIDSQQEQIYIQCIEKLIEGCFEGYNATV
1005						FAYGQT\GAGKTYTMGTGFD
1096	2446	Α	9029	1	285	FFFFNVCKSPKVPKPGCKEESTGTLFKNTLISL
						GQHSETPSLKKK\LAGYSGMCL*SQVLRRLRQ
						EDCLSPGGGNCRES*SCPYTPAWITERDPV
1097	2447	Α	9032	716	357	ARSTGFWGEILWCGFLKRSLALSPRVKCSGAI
	·					LAHCNFRHAGFPPLSCLSLPNRWEYRRPPARP
		ł	}	-		GKFFLVFLVETGFQC/G*DGLDLLTSRSACLG
						LPKCWDYRREPAASIIFQTTFFINSK
1098	2448	Α	9038	230	652	KVVVMSCEDINISGSFYRNKLKYLAFLCKRTS
	1	İ				TNPSQGPYHLWVPSHIFWQTTCGRLPHKTKO
1 1	]	J		j		G*AALDHLKVFDRIPLPYDKKKQMAVSATLE
	ľ	1	l			VVRPKP*RKFAYLGHWAQKVDWKYQAMTA
		l			-	TMGEKRKVYYQKICYQKK
1099	2449	Α	9043	185	372	IIFYSHQQCMRV/WQGCGDIETLIHCW*E*KII
	J			i		HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP
						RKIKTCPQNSCTSMLINAIHNDQKWKKINI
1100	2450	Α	9045	763	584	RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL
					1	SLTSSWDYRRPPPHPANFLYFK*RRGF
1101	2451	Α	9050	275	2	LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL
	į	- 1	- 1	ł	1	FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS
		1		l.	1	DSPASAS*VAGITDMCRYTQLILFHAS
1102	2452	A	9053	449	1224	KTSMFWKFDLHSSSHIDTLLEREDVTLKELM
	ŀ	ĺ	[	ĺ	i	DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF
				I		N*EEPPQDMDEKIRYKYPNISCELLTSDVSQM
	i	i	1			NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF
[	{	[	1		ĺ	SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG
		İ	1	ł	J	TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ
		ľ			İ	RNL*HST*NVMDISKYVNLHWGLNKSHSLL*
						LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS
·		ł	1	f	ł	SLV
1103	2453	A	9058	403	3	GLHVYDFQVYREHILTLNVKKCSVSFWGLRE
-					-	WLYLQMYEIIKSPRFPIIKMTDITKCW+GC\GA
i	ĺ	- 1	1		1	AGMQLH/CW\WCVNVGKFWEMS*YYLLKLSI
Į		Ţ	1	1		ST/PYDPAIPLLGIYL*ETRYYIHPKTCMRMLIA
						THE TOTAL BUTTE BIRTINIA

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nuci-	peptide	поп	in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
		[		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ	İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ł	ł	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						APFVLAVNC
1104	2454	Α	9064	75	393	KWLFSSLNITGRGDIIGHLKWLDCR\NCSSFPI
						KRNRQTHSTESNKLKAGHSFGYN*LIH*NS\V
		i				KTDCGCGANSKGVVVVMKV\KTAQQKQTTS
l						YMQIGTTKNSRAT
1105	2455	A	9065	366	778	DLLILRNLAFPELKRRNCISRFYLAYHLHKIYS
1				i		RSILLCNNCSGFYILSL*QYDVFFFNYFFFRDR
1			Ì			AWPCCPGWSAAWLTTVILAHYRRPGLERSCC
1						LSLSSSWDHRRVPPCPANF*/YFSMGFTAFPRL
						VLNS*TQGI
1106	2456	A	9083	673	816	ESGSLIH*WWENKPAQPLWWEI*QHVQKLPT
1100	0455		0000			HFPCDPAIPLLGICPED
1107	2457	A	9086	580	18	KPSSGSFIRAJYIFLSTAHVPALFSVLVRTKLT*
						AFSQSSVLWAHKQQKTSLSLVIR/ERLQIKTA
						VRENFLPIRLAKILKLDNVKCWQG/SGSNMSL
1						UHCWWEYNVIHIIWNSVTFPRKVEHVYITYA
]						PEISVR*IHGGLPTLVHQETHTSVFRGAPSVIP
1108	2458	A	9093	540		ETR\CRPTKESINKLLHIYTMEHYGDENK
1100	2430	A	5093	340	1	GGNDCSVTPTTEPGRKEIT*KRKF*EKTDRLP
						GA/PPSRTPPTPYPCPHGDRLLPPSRPLPAGPA
						SAFPPAERSRGHRRASL*RARWSAAVPRRSA
						GSASEPVQSRWLRLPVGSDSPPAVPVRVCPAP DSRPAAPGSRLPDPGLDSPAPSRTPSSSVD*GG
Ì						QRPPPPSGDSLSPPGCCRY
1109	2459	A	9099	1255	1425	HESYHVNPNLCNPVAPTSGAHSIG*KWPSWL
1	,		3033	1200	1123	GAVAHSCNPSTLVGRGGRITRGQELR
1110	2460	A	9103	242	70	EEQFFFFAVGMFP*VDFLAPASGELWDRLRLT
						CSRPFTRHQSFGLAFLRVCSSLDSLDDSVVGP
						SALLSSVL/NQGGRNVLEAREAAKHPTI*RQS
						LLRKQRNKRMAIP
1111	2461	Α	9110	189	121	SFLSVRLECNGAIMAHCALPLPG
1112	2462	A	9113	100	910	RRRGGGSRPRRTPVPAPGPGPSFGMDVRFYP
1						AAAGDPASLDFAQCLGYYGYSKFGNNNNYM
1						NMAEANNAFFAASEQTFHTPSLGDEEFEIPPIT
}						PPPESDPALGMPDVLLPFQALSDPLPSQGSEFT
<u> </u>				i		PQFPPQSLDLPSITISRNLVEQDGVLHSSGLHM
				l		DQSHTQVSQYRQDPSLIMR\PSST*PDAARSG
						VMPPAQLTTINQSQLSAQLGLNLGGASMPHT
						SPSPPASKSATPSPSSSINEEDADEANRAIGEK
1113	2463	A	9120	3452	2051	RAAPDSGKKPKTPKK
1113	2403	n.	7120	3432	3051	FLRPSFALVPQAGVQWCALSWLQPPSPRFK*F
					i	SCLSLPSSWDYRHVPPRPANFFVLLVETGFLH
						VGQAGHEPLTSGDPPASASQSAGITGVSHQA WPSFFIFSRDTVLLCCSGWSRTSGLKQSACLS
( [						
1114	2464	A	9122	152	377	NOLPLOQWIFFIYETGFCSVAQAGVQCRDHS
····	2.07	**	7100	.52	3,,	SLHP*PPG\SSDPPAPPS*VLGITGQRYHACLII
				ļ	}	YLYVQTVPQRV
1115	2465	Ā	9124	553	981	QRPLLRQQLGSWPTCRSLEGDLASPW**RLPG
	55		7.27		701	SPRMRRSGT/ATLNLPLSPQGTVRTAVEFQVM
]				ļ	1	TQTQSLSFLLGSSASLDCGFSMAPGLDLISVE
ļ [	1	- 1	[			WRLQHKGRGRGDLHLPDHHLSVPSSADHPA
	,					QQPSQFNGRNLYFLPLFR
1116	2466	A	9135	48 .	410	SASHEPAEHDGGADSLSASQPPRPAGRPAGA
						QHVHVPPWTDVLAGQDRRAPTAGDGAPWP
	į	ł				APGGHVPSTRPHDPAEFHADEAAGRGGRGLQ
	į					PAAPHALPAGLPHGPPAPA/PAEGGGTP*GSA
					İ	GAGGP*GSPAGRACGAAGCRPRPPRPAASSA
	-				j	*NSAGS*GLVEGT*PPGAGHGAPSPAVGARLS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  CPARTSVQGGTWTC*APAGRPAGLGGWEAE RESAPPSCSAGS*DAD*GAEPWGAGSRSWGS
		^		380	737	KSGHWAKECLQPRIPPRPCPICVGPHWKSDCP TCPGAVPRAPGTLPQGSLTDSFPDILSLVAED *CCLMASEASWTIT\ELWVTLTVEGKSVP/CL NTEATHSTLPSFQGPVSLASITVVGIDGQASKP LKTPQLWCQLGQYSFMHYFLVIPTCPVPLLG* GILTKLSAFLTIPRLQPHLIAALSPSS
1118	2468	A	9154	471	2	AAGQVVVEVTSHLYLCITSDAAGLRILPPAES ERGEGGHCPAEAPLPPRPQYCLAKHPLLRKLP EEKIKLDPYLTQHTKINSKQIKYLS/VRAKTTQ LVEGNIGVNLQNTELKQH*INGFLDTTPEAQE TKEKTNKLNFIKKVKRQLAEWEKIFQIA
1119	2469	A	9155	2	3187	ACPRLARRRRRVRSLRRRRGWLRARWSRGQ NNMAARRITQETFDAVLQEKAKRYHMDASG EAVSETLQFKAQDLLRAVPRSRAEMYDDVHS DGRYSLSGSVAHSRDAGRESLRSDVFSGPSFR SSNPSISDDSYFRKECGRDLEFSHSNSRDQVIG HRKLGHFRSQDWKFALRGSWEQDFGHPVSQ ESSWSQEYSFGPSAVLGDFGSSRLEKECLEK ESRDYDVDHPGEADSV/LRGGSQVQARGRAL NIVDQEGSLLGKGETQGLLTAKGGVGKLVTL RNVSTKKIPTVNRITPKTQGTNQIQKNTPSPD VTLGTNPGTEDIQFPIQKIPLGLDLKNLRLPRR KMSFDIIDKSDVFSRFGIEIIKWAGFHTIKDDIK FSQLFQTLFELETETCAKMLASFKCSLKPEHR DFCFFTIKFLKHSALKTPRVDNEFLNMLLDKG AVKTKNCFFEIIKPFDKYIMRLQDRLLKSVTP LLMACNAYELSVKMKTLSNPLDLALALETTN SLCRKSLALLGQTFSLASSFRQEKIL*AVGLQ DIAPSPAAFPNFEDSTLFGREYIDHLKAWLVS SGCPLQVKKAEPEPMREEEKMIPPTKPEIQAK APSSLSDAVPQRADHRVVGTIDQLVKRVIEGS LSPKERTLLKEDPAYWFLSDENSLEYKYYKL KLAEMQRMSENLRGADQKPTSADCAVRAML YSRAVRNLKKKLLPWQRRGLLRAQGLRG\ WKARRAITTGTQTLLFLRAPGLKHHGRQAPG LSQAKPSLPDRNDAAKDCPPDPVGPSPQDPSL EASGPSFKPAGVDISEAPQTSSPCPSADIDMKT METAEKLARFVAQVGPEIEQFSIENSTDNPDL WFLHDQNSSAFKFYRKKVFELCPSICFTSSPH NLHTGGGDTTGSQESPVDLMEGEAEFEDEPP PREAELESPBVMPEEEDEDDDGGEEAPAPG GAGKSEGSTPADGLPGEAAEDDLAGAPALSQ ASSGTCFFRKRISSKSLKVGMIPAPKRVCLIQE PKGECPPVGTVASSTVLGWWAVRVRRDRWR HFNPKEFCAPLQNVSRHSCFPVV
1120 1121	2470 2471	A	9163 9166	124 272	207	PPRACRPCPRACPCPPT*KCSQPVSWPC
1122					523	PMSSLQGCFYTFKCIIFKGIFLLLISNLIAF**EK V/CSHITDSLKFIGKGWVGMVTHACNPGTLG G*GGWIA*VREFETSLGNM
			9170	442		MNRRFLRPADCHSGMRGTENGACSEGESQI HCGAGGEGVQLVHVVNQPENGCLQFDSTHIT FSKRQN*
1123	2473	A	9171	10		MVDRSPLLTSVIIFYLAIGAAIFEVLEEPHWKE AKKNYYTQKLHLLKEFPCLGQEGLDKILEVV SDAAGQGVAITGNQTFNNWNWPNAMIFAAT VITTIGYGNVASKTPGGRLFCGFYGLFGVPFC LTWINALGKFFG

SEQ ID	SEQ ID	Met	SEQ	Predicted	18 0.1	
NO: of	NO: of	hod	ID NO:	beginning	Predicted and	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1104	in NO.	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		USSN	location		F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence	i	09/496	correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1		1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	}	ł	]	peptide	l sodemen	/=possible nucleotide deletion, \=possible
1		1		sequence		nucleotide insertion
1124	2474	A	9173	3	374	GPSPSLLVLLPQEPGGTGTPVRAGAGAGMWL
1		i		1		WEDQGGLLGPFSFLMLMLLLETRNPVNACLL
1	ŀ				1	TGSLFVLLGVFSFEPVPSCRALQELKPRDRISA
L		1	i			IAHRGGRHDPPENTLGAIR/QGS++WSNRR
1125	2475	A	9179	704	188	ESSSGLLFQCFQGIHVQKLTLQARPTLFSWWL
	}				-	CSKPPKETGELENAESGGDGGRRGGKQDNV
1	ļ		1		·	AWWRRM\QKG\DFPWDDEDFPQSGPFGGQA
		İ	Ì			LPMGFFYLYFRDPGREITWKHFVQYYLARGL
1	Į	ł	1			VDRLEVVNKQSVRVIPAPGTSSEVRGEFKAE
			<u> </u>			YCRHKFISCKNVVFYFFQ
1126	2476	A	9183	153	233	MEYMAESTDRSPGHILCCECGVPISPN
1127	2477	Α	9185	1	321	LTGQLGSILLRVFSKSRAGLGARKLKAYRTM
				·		EYMAESTDRSPGHILCCECGVPISPNPAQY\CV
1		١.				ACLRSSFHIYHCIPKLFIHPFSKTSSSAFITPSHY
1100	0.450	<u> </u>				LTFFSTIS
1128	2478	A	9186	183	847	VLKFLLLQTMDEQSQGMQGPPVPQFQPQKAL
1						RPDMGYNTLANFRIEKKIGRGQ\FSEVYRAAC
	j :					L\LDGVPVALKKVQIFDLMDAKARADCIKEID
			İ			LLKQLNHPNVIKYYASFIEDNELNIVLELADA
					,	GDLSRMIKHFKKQKRLIPERTVWKYFVQLCS
			·			ALEHMHSRRVMHRDIKPANVFITATGVVKLG
1	1 1					DLGLGRFFSSKTTAAHSLVGTPYYMSPERIHD NG
1129	2479	Α	9190	1	370	GTSWKIPSAAVSESSPNGAAYASGLPCGVRG
1			71,70	•	370	PPWAGLALLPSPTLMALLRRPTVSSDLDNIDT
						RATTIKIRVVATTTRARIEDMRHSATALTRPD
				- 1		ATTAQIPKLPVTTVCNRRANPGIPPSVL
1130	2480	A	9194	131	487	AYLKRLPVPESITGFARLTVSEWLRLLPFLGV
						LALLGYLAVRPFLPKKKQQKDSLINLKIQKEN
1				ł		PKVVNEINIEDLCLTKAAYCRCWRSKTFPAC
						DGSHNKHNELTGDNVGPLILKKKE
1131	2481	Α	9201	184	605	KELVDEKSERGRAMDPVSQLASAGTFRVLKE
]			j	ļ	ļ	PLAFLRALELLFAIFAFATCGGYSGGLRLSVD
				İ		CVNKTESNLSIDIAFAYPFRLHQVTFEG\PTCE
		,				GKERHKLALIGDSSSSAEFFGTVAGFAFLYSL
1130	2400		-0055			AATGVYIFFQNKY
1132	2482	A	9206	1	852	GGGRAGAGSRDMGSTDSKLNFRKAVIQLTTK
				ł	. –	TQPVEATDDAFWDQFWADTATSVQDVFALV
		ļ		}		PAAEIRAVREESPSNLATLCYKAVEKLVQGA
j l			ĺ	<b>!</b>		ESGCHSEKEKQIVLNCSRLLTRVLPYIFEDPD
				ľ	1	WRGFFWSTVPGAGRGGQGEEDDEHARPLAE
			<b>[</b>			SLILAIADLLFCPDFTVQSHRRSTVDSAEDVH
	i .		l			SLDSCEYIWEAGVGFAHSPQPNYIHDMNRME
; l	<b> </b>		ļ	1	- 1	LLKLLLTCFSEAMYLPPAPESWQH/RTHWFSS
	i	ļ	ŀ	l		FVSSENRHALPLFTSLLNTVCAYDPVEYGIPY NHLY
1133	2483	A	9208	1165	1463	
			/200		1400	GPRARVQGFSGADIVKFMALGSMYLVLTLIV
1	ı		- 1	i	ł	AKVLRGAEPCCGPLKNRVLRPCPLP/VPLPPP
					İ	HPQPSRGNPVGCLPTYKVVYKLLSWPLHSNS NVYFIV
1134	2484	A	9210	66	1586	MAGAGPKRRALSAPVAEEKEEAREKIMAAK
		J			1300	RADGAAPAGEGEGVTLQGNITLLKGVAVIVV
	į		- 1	}	l	AIMGSGIFVTPTGVLKEAGSPGLALVVWAAC
		1		Ì	1	GVFSIVGALCYAELGTTISKSGGDYAYMLDV
	I	ŀ	i		i	YGSLPAFLKLWIELLIIRPSSQYIVALVFATYL
	- 1	ł	- 1	1	ļ	LKPLFPTCPVPEEAAKLVACLCVLLLTAVNC
						YSVKAATRVQDAFAAAKLLALIILLGFVQ1
						GKGDVSNLDPNFSFEGTKLDVGNIVLALYSG
	}		I		j	LFAYGGWNYLNFVTEEMINPYRNLPLAIIISLP

NO: of nucle order of the property of the control		SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Section   Sect				hod	ID NO:		nucleotide	D=Aspartic Acid, E=Glutamic Acid,
1136	١							
1914   mg to first amino acid residue of peptide sequence   msidue e   msidue of peptide   msidue	ı							
	Ī	-	uence				, ,	
Personal	1	HEHCE			914			
Popsible nuclotide deletion, 1-possible nuclotide insertion   Popsible nuclotide insertion   Popsible nuclotide insertion   Popsible nuclotide insertion   Popsible nuclotide insertion   Popsible nuclotide insertion   Popsible nuclotide insertion   Popsible nuclotide insertion   Popsible nuclotide insertion   Popsible nuclotide insertion   Popsible nuclotic nuclotide   Popsible nuclotic nuclotide   Popsible nuclotide nuclotide   Popsible nuclotide nuclotide   Popsible nuclotide nuclotide   Popsible nuclotide nucloti	ı							
	ı						Suppose	=nossible nucleotide deletion \=nossible
INTLUVYLINLAYFITLSTEQMLSSEAVAND	ı							
SNYHLGYMSWIPPVOLSCFGSVNGSLTPSLTYT   CVMTLFYARSKDIFSVNFRSFFNWLCVALAB   GMWM.BRIRKPELBRIMBIPQL_LITPYSLVFT   CVMTLFYARSKDIFSVNFRSFFNWLCVALAB   GMWM.BRIRKPELBRIKNMLEQL_LITPYFILACLF   LIAVSFWKTTFWSVASDFTILLSGLPVYFPG   WWKNKPKWAPGWAPGHLSPRRSVCRSCMVYPQ   WWKNKPKWAPGMLSPRSVCRSSCMVYPQ   WWKNKPKWAPGMLSPRSVCRSSCMVYPQ   WWKNKPKWAPGMLSPRSVCRSSCMVYPQ   WWKNKPKWAPGMLSPRSVCRSSCMVYPQ   WWKNKPKWAPGMLSPRSVCRSSCMVYPQ   WWKNKPKWAPGMLSPRSVCRSSCMVYPQ   WWKNKPKWAPGMLSPRSVCRSSCMVYPQ   WWKNKPKWAPGMLSPRSVTSSCMVYPQ   WWKNKPKWAPGMLSPRSVTSSCMVYPQ   WWKNKPKWAPGMLSPRSVTSSCMVYPQ   WWKNKPKWAPGMLSPRSVTSPRSVTSPRPCPLAVSMPHAFKPQ   DLVFARMKOPFWPARAIDIADOAVXPPTAFKPQ   DLVFARMKOPFWPARAIDIADOAVXPPTAFKPQ   PNRRKCFPRGLWEIQNNPHASYSAPPVSSSS   SRAPFANPADOSDADEDBGRGWAPAVTAVT   ATAASDRMEDDSDKSSONSGLKKKTTAK   MSVSKRARKASDADEDBGRGWAPAVTAVT   ATAASDRMEDDSDKSSONSGLKKKTTAK   MSVSKRARKASDADEDBGRGWAPAVTAVT   ATAASDRMEDDSDKSSONSGLKKKTTAK   MSVSKRARKASDADEDBGRGWAPAVTAVT   ATAASDRMEDDSDKSSONSGLKKKTTAK   MSVSKRARKASDADEDBGRGWAPAVTAVT   ATAASDRMEDDSDKSSONSGLKKKTTAK   MSVSKRARKASDADEDBGRGWAPAVTAVT   ATAASDRMEDDSDKSSONSGLKKKTTAK   MSVSKRARKASDADEDBGRGWAPAVTAVT   ATAASDRDATKAT   ATAASDRMEDDSDKSSONSGLKKTTAK   MSVSKRARKASDADEDBGRGWAPAVTAVT   ATAASDRDATKAT   ATAASDRMEDDSDKSSONSGLKKTTAK   MSVSKRARKASDBGARPPTAWARSASSSS SSSSDSDSVVKKPPRGRKPAEKPIP,FKRGRGWGRGKK   RGSREPPAWA   STORT   MSVSWATCHTAKT   MSVSWA	Γ					-		IVTLVYVLTNLAYFTTLSTEQMLSSEAVAVDF
CVMILFYARSKOJESVINFESFENWI.CVALAID	1							GNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSS
1135   2485   A   9216   40   410   410   RDRLPFRYCTURE WWKNIKEWAPPGHLSPRSCOMPYPG WWKNIKEWAPPGHLSPRSCOMSCOMPYPG WWKNIKEWAPPGHLSPRSCOMSCOMPYPG WWKNIKEWAPPGHLSPRSCOMSCOMPYPG WWKNIKEWAPPGHLSPRSCOMSCOMPYPG WWKNIKEWAPPGHLSPRSCOMSCOMPYPG RDRLPFAYFCRPVVCVVTALDVGSSENGMYPG DLVAFEDVANNTTQEEWSLLDFSQKNLTREW MQETLRILASIGEKWKDQNIEDQYKNPRING RSLEGRYDDTINK RSLLGERVDENTERNINGGISTSQUPDTINK RSLEGRYDDTINK RSLEGRYDDTINK RSLEGRYDDTINK RSLEGRYDDTINK RSLEGRYDDTINK RSLEGRYDDTHERSHINGGISTSQUPDTINK RSLEGRYDDTHERSHINGGISTSQUPDTINK RSLEGRYDDTHERSHINGGISTSQUPDTINK RYPIFFGTHETAFLGYKDLFPYDKCKDKYG PKVRIFFATHETAFLGYKDLFPYDKCKDKYG PKVRIFFATHETAFLGYKDLFPYDKCKDKYG PKVRIFFATHETAFLGYKDLFPYDKCKDKYG PKVRIFFATHETAFLGYKDLFPYDKCKDKYG PKVRIFFATHETAFLGYKDLFPYDKCKDKYG RYSKKARKASSDLDQASVSPSEBENSESSE SEKTSDDDTFTEKKAAVRAPRGPFYSSLENSESSE SEKTSDDDTFTEKKAAVRAPRGPFYSSLENSESSE SEKTSDDDTFTEKKAAVRAPRGPFYSSLENSESSE SEKTSDDDTFTEKKAAVRAPRGFTFKGK RYSKETTAFLGYKDLFPYDKCKDKYG GSTWNTRGGLPHTNTFYKTGFTCSQAGLKL RGSREPPAWA RSASSSSSSSSDSDVVKKPRGRKPAEKPLPKTGRK RKFTAKKAPSSLESSESSESSDSDVVKKPRGRKPAEKPLPKTGRK RKFTAKKAPSSLESSESSDSDVVKKPRGRKPAEKPLPKTGRGGRGK RGSREPPAWA RSASSSSSSSSSSDSDVVKKPRGRKPAEKPLPKTGRGGRGK RGSREPPAWA RSASSSSSSSSSDSDVVKKPRGRKPAEKPLPKTGRGGRGK RGSREPPAWA RSASSSSSSSSSSDSDVVKKPRGRKPAEKPLPKTGRGGRGK RGSREPPAWA RSASSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	1							
LIAVSFWKTTPWSVASDFTIILSGLFVYFFGV	L							CVMTLFYAFSKDIFSVINFFSFFNWLCVALAII
WWKNKYEWAPPGHLSPRPSCYRSSCMVYPQ	l						ļ ,	)
1135								1
DLVAFEDVANFITOEBWSLLDPSQENLYTES	r	1135	2485	Α	9216	40	410	
MQETLRNLASIGEKWKDQNIEDQYKNPRINNI   1136				77		10		DLVAFEDVAVNFTOEEWSLLDPSOKNLYREV
RSLLGER/DENTENHGEITS/GEDPOTIANK	L							MQETLRNLASIGEKWKDONIEDOYKNPRNNL
DLVFAKMKGYPHWPARIDDIADGAVKPPPN KYPIFFGTHETA-I-GYRDLPFYDKCDK-YGE PNKRKGFNEGL-WEIQNNPHASYSAPPPVSSSI SBAPPANPADGSDADEDDEGRGVMAVTAAVI ATAASDRMESDBSDANSSSDNSGLKRKTPALK MSVSKRARKASSDLDQASVSPSEENISSSSS SEKTSDQDFTPEKKAAVASPRRGPLGGRKKK APSASDSDXADSDAKPEPVAMARSASSSS SSSSDDVSVKKPPRGRKPAFKPLPKPRGRK KPERPPSSSSD UFFRLECRDPVTVNCTLNLPGSKNAFTTASQ STWNYRGGLPHTMPFVKTGFRCSQAGLKL RGSRPPAWA  1138 2488 A 9231 1664 2 TRSVGVNTCEGUVVTEPECLGPCEPGTSVNL BGIVWHETEEGVLVVNVTWRNKTYVGTLLD CTKHDWAPPRFCESPTSDLEMRGGRGRGK ARSAAAAAPGSEASFTESRGLQKKNRGGANGK GRGSLNASGRTPPNCAAEDIKASPSTINKR KKPPMELDLINSSEDNKGKRVRTNSRST TTPQGKPETTFLDQGCSSPVLDENKIRFGPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAONPPGTPKGKRELMSNOPGSIIGAKAGK VKHINGLRYHQAHAHLDKSSCANADGSSPVLDENKIRFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAONPPGTPKGKRELMSNOPGSIIGAKAGK ANNCKIDKNPSKLKSKRIMSNOPGSIIGAKAGK ANNCKIDKNPSKLKSKRIMSNOPGSIIGAKAGK KKLKNEPMELDKHLNNSESIDLANGLISEGSERMAS ANNCKIDKNPSKLKSKRIMSNOPGSIIGAKAGK KKLKDRPSKLKSRIMSHOPGSIIGAKAGK KKLKDRPSKLKSRIMSHORGSIIGAKAGK LIKDKEGKETGSPKMDAKLGKLEDSKGASK DLEPHFIKINGEPTIVPHAPAPTPPPQLIA PTAITFITTTTGTIPGLPSLTTTVVQATPKSPPL KPIOPKPTIMGEPTIVPHAPAPTPPPQLIA PTAITFITTTTGFIPLPSLTTTVVQATPKSPPL KPIOPKPTIMGEPTIVPHAPAPAPTPPQLIA PTAITFITTTTGFIPLPSLTTTVVQATPKSPPL KPIOPKPTIMGEPTIVPHAPAPAPTPPQLIA PTAITFITTTTGFIPLSLANGLISESGESRMAS LLEPHFIKINGHINNERLISERGALIGISESGESRMAS IKAEADKVYTFIDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGGPWRRRAAAAGILPGREAAACIPSC/AS TRRGGPWRRRAAAAGILPGREAAACIPSC/AS TRRGGPWRRRAAAAGILPGREAAACIPSC/AS TRRGGPWRRRAAAAGILYGREAAACIPSC/AS TRRGGPWRRAAAAAGILYGREAAACIPSC/AS TRRGGPWRRAAAAGILVGYEGISGALLQKWILLALSC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGGTSNGADESPNMLVYRKY 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRKALAFRILLRILERSP TLRGHGGASGRNVTTGSLGEPGWRVATGG RPGTSPALLFSGRGAATGGQGGGFDTKCLAA ATWGRLPOPEETLFQQDSWNCVPRACILGHA WPWAAALVVHCYSKSSPSNKDAALLEAARAQ NMGGVSRNCCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFYGFLFEPGQDSWNCVPRACILGHAACQCC PALGNDAIKTELLTMAGQCC PALGNDAIKTELLTMAGQCC PALGNDAIKTELLTMAGQCC FALGNDAIKTELLTMAGQCC FALGNDAIKTELLTMAGQCC FALGNDAIKTELLTMAGGCC F	L							
KYPIFFGTHETAFLGPKDLFPYDKCKDKYOK   PNRKKGPNEGLWEIDNPHASYSAPPPVSSS   SEAPEANPADGSDADEDGIGGVMAVTAVT   ATAASDRMEBDSDSDKSSDNSGLKRKTPALK   MSVSKRAKASSDLDQASVSPSEEERSSSSE   SEKTSDQDFTPEKKAAVRAPRGPLGGRKKK   APSASDSSKADSDGAKPPVAMARSASSSS   SSSSSDDVSVKKPPRGRKPAEKPLPKPRGRK   PKPERPPSSSSD   LFPRLECRDPVTVNCTLNLPGSKNAPTTASQV   GSTWNYRGGLPHTNFPVKTGFRCSQAGLKL   RGSREPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPPAWA   RGSPAWA		1136	2486	Α	9223	3	983	RRRRSRYRRCSRFPRPGPLAVSMPHAFKPG
PNIKRKGFNEGLWEIQNNPHASYSAPPPVSSSI   SEAPEANPADGSDADEDGGRGVMAVTAVT   ATAASDRMESDSDSDSSDSSSNOGLKRKTPALK   MSVSKRARKASSDLDQASVSPSEENSSSSSSSSSDSGLKRKTPALK   MSVSKRARKASSDLDQASVSPSEENSSSSSSSSSSSSDSDVSVKKPPRGRKK   APSASDSDSKADSDGAKPEPVAMARSASSSSS   SSSSSDSDSVSVKKPPRGRKKPPKPRGRK   PKPERPPSSSSD   LFPRLECRDPYTVNCTLINLPGSKNAPTTASQV   GSTWNYRGGLPHPTNFYVKTGFRCSQAGLKL   RGSREPPAWA   RGSREPPAWA   RESPERA   PROPERTY   RESPECTIVE	t					-		
SEAPEANPADGSDADEDDEGRIGVMANTATY   ATAASDRMESDISDSSSNSGLKKRTPALK     MSVSKRARKASDLDQASVSPSEEENSESSE   SEKTSDQDFTPEKKAAVRAPRGPIJGGRKKK     APSADSDSKADSDGAKPEPVAMARSASSSS   SSSSDDVSVKKPPRGRPPAERPLFQRGK     APSADSDSVADSDGAKPEPVAMARSASSSS   SSSSDDVSVKKPPRGRKPAEKPLPKPRGRK     FYPERPSSSSD   LFPRLECRDPTVNCTLNLPGSKNAPTTASQV     GSTWNYRGGLPHPTNFFVKTGRCSQAGLKL     I138   2488   A   9231   1664   2   TESVGVNTCEVGVVTEPECLGPEGTSVN     LGGVWHETEEGULVNVTWRNKTYVGTLLD     CTKHDWAPPRFCESPTSDLEMRGGRGRGKR     ARSAAAAPOSEASPTSDLEMRGGRGRGKR     ARSAAAAPOSEASPTSDLEMRGGRGRGKR     ARSAAAAPOSEASPTSSULDPHPNCNKK     YKHINGLRYHQAHAHLDPENNKLEFEPDSEDK     STORGELSNVALECSPTSVALPPINCNKK     YKHINGLRYHQAHAHLDPENNKLEFEPDSEDK     STORGELSNVALECSPTSVLADPHONKERGK     ANNCKTDKNVSKLKSARPIAPPAPPTPPQLLA     BPTATFTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	l							
ATAASDRMESDSDSSKSSDNSGLKRETPALK   MSVSKRARKASSDLDQASVSPSEENSESSSE   SEKTSDQDFTPEKKAV VRAPRGPLGGRKKK   APSASDSDSKADSDAKPPV AMARSASSSS   SSSSDSVVSVKPPRGRKPAEKPLYRPGGRK   PKPERPPSSSSD   LFPRLECRDFVTVNCTLNLPGSKNAPTTASQV   GSTWNYRGGLPHPTNFFVKTGPCSQAGLKL   RGSREPPAWA   RGSREPAWA   RGSREPAWA   RGSREPAWA   RGSREPAWA   RGSREPAWA   RGSREPAWA   RGSREPAWA   RGSREPAWA   RGSREPAWA   RGSREPAWA   RGSREPAWA	L							
MSVSKRARKASDILDQASVSPSEEENSESSES   SEKTSDQDFTPEKKAAVRAPRAPRGPIGGRKKK APSASDSDSKADSDGAKPEPV AMARSASSSS: SSSSSDSDVKKPPRGRKPAEKPLPKPRGRK PKPERPPSSSSSD    1137   2487   A   9229   21   239   LFPRLECRDPVTVNCTLNLPGSKNAPTTASQV GSTWNYRGGLPHPTNFFVKTGPRCSQAGLKL RGSREPPAWA     1138   2488   A   9231   1664   2   TRSVGVNTCEVGVVTEPECLGPCEPGTSVNL EGIVWHETEGALVANVTVPRNKTYVGTILD CTKHDWAPPRFCESPTSDLEMRGGRGRGKR ARSAAAAPGSEASFTESRGLNKNRGGANOK GRROSLNASGRTPPNCAAEDIKASPSSTNKR KNRFPMELDLNSSSEDNKPGRRVRTNSRSTP TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAVDQLKAPA SPGAGNPGTPKGKELMSNGPGSIGMAAGK NSGKKGLNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDICKKKKKK NKSGKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKNLGDICKKKKKKK NKSGKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKNLGDICKKKKKK KKKDKEPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTFDGLPSLTTTVQATPKSPPL KPIQKPTIMGEPINPALVSILKDKKKKKKK KLKDKEGKETGSPKMDAKLGKLEDSKGASK NSGKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKNLGDICKKKKKK KLKDKEGKETGSPKMDAKLGKLEDSKGASK NSGKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAGGLIDKNLGDICKKKKKK KLKDKEGKETGSPKMDAKLGKLEDSKGASK NSGKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAGGLIDKNLGDICKKGKK KLKDKEGKETGSPKMDAKLGKLEDSKGASK NSGKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAGGLIDKKKKEKEK KLKDKEGKETGSPKMDAKLGKLEDSKGASK NSGKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAGGLIDKIKTLGLESGESGESMAS KAEADKVYTTINPAPLSISLGTGAGATGROVENG TARGGPWGNSLVIGALL SC HEQEMGVSSLVIGALL    1140								
SEKTSDQDFTPEKKAAVRAPRGPI,GGRKKK APSASDSDSKADSDAKPEPVAMARSASSSS SSSSSDSDVSVKKPPRGRKPAEKPLFKPRGRK PKPERPPSSSSSD  1137 2487 A 9229 21 239 LFPELECRDPVTVNCTLNLPGSKNAPTTASQV GSTWNYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA  1138 2488 A 9231 1664 2 TRSVGVNTCEVGVVTEPECLGPCEPGTSVNL EGIVWHETEGVLVVNVTWRNKTYVGTLLD CTKHDWAPPRFCESPTSDLEMRGGRGRGKR ARSAAAAPGSEASFTESRGLQNKNRGGANOK GRRGSLNASGRTFPNCAAEDIKASPSSTNKR KNKPPMELDLNSSSEDNKPGKRVRTNSRSTP TITPQGKPETTFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPFNKLEFEPDSEDK ISDCEEGLSINVALECSEPSTSVSAVDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNBLINNLPVISNMTAALDSCSAA DGSLAAEMPKLEAGLIDKKNIGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFITTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMEDEJTVNPALVSLKDKKKKEKK KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHLKDHLINKNEGLANGLESSCGSRMAS IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQFWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGGTSNGVADESPNMLVYRKV LEQEMGVSSLVIGALL  1141 2491 A 9242 2 535 FYEAAVKMIGSLVLRKKALAPRILLRLEXP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGRGGGRTDTKCLAA ATWGRIPGPETLIPGQDSWNGVPSRAGLGMA WPWAAALVYHCYSKSPSNDAALLEAARAQ WPWAALLVYHCYSKSPSNDAALLEAARAQ WPWAALLVYHCYSKSPSNDAALLEAARAQ WPWAALLVYHCYSKSPSNDAALLEAARAQ WPWAALLVYHCYSKSPSNDAALLEAARAQ WPWAALVYHCYSKSPSNDAALLEAARAQ WPWAALLVHCYSKSPSNDAALLEAARAQ WPWAALLVHCYSK			Į					
APSASDISKADSIGAKPEPVAMARSASSSS   SSSSISDISVKKPPRGKKPAKPLPKPRGKK    PKPERPPSSSSD    FYPELECRIPYTVICTILIPGSKNAPTTASQV    GSTWNYRGGLPHTINFFVKTGFRCSQAGLKL    RGSREPPAWA	1		1					I =
PKPERPPSSSSSD		-						<b>APSASDSDSKADSDGAKPEPVAMARSASSSSS</b>
1137   2487   A   9229   21   239		.	1					SSSSSDSDVSVKKPPRGRKPAEKPLPKPRGRK
STWNYRGILPHPTNFFVKTGFRCSQAGLKL   RGSREPPAWA	$\vdash$	1127	2407	<del>_</del>	2000		000	
RGSREPPAWA	1	113/	2487	Α	9229	21	239	
1138								
EGIVWHETEEGVLVVNVTWRNKTYVGTLLD CTKHDWAPPRFCESPTSDLEMRGGRGRGKR ARSAAAAPGSEASFTESRGLQNKNRGGANGK GRRGSLNASGRRTPPNCAAEDIKASPSSTNKR KNKPPMELDLNSSSEDNKPGKRVRTNSRSTP TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPPPPQLIA PPTATFTTTTGTPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRRAAAAGILFGREAAACLPSCIAS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAFRLLLRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRPDTKCLAA ATWGRLPOPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NNGEVSRRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFFIKKTC CTSMDAIEKTEPLTKWTKKAFVSHVSYKRLLF	r	1138	2488	A	9231	1664	2	
CTKHDWAPPRFCESPTSDLEMRGGRGRGKR ARSAAAAPGBASSTFESRGLQNKNRGGANGK GRRGSLNASGRTFPPNCAAEDIKASPSSTNKR KNKPPMELDLNSSSEDNKPGKRVRTNSRSTP TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPPTPPQLLA IPTATFTTTTTGTPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKGKEGTGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKYVTTIONAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTILLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRPDTKCLAA ATWGRLPGFEETLPGQDSWNGVPSRAGLGM WPWAAALVYHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQATKEKLDKLDFIKIKTC CTISMDABEKTEPLTKWTKKAFVSHVSYKRLLF							_	
GRRGSLNASGRRTPPNCAAEDIKASPSSTNKR KNKPPMELDLNSSSEDNKPGKRVRTNSRSTIP TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIGAKAGK NSGKKKGLNNELNNLPVISIMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKER KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFIKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFIDNAPSPSIGS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGYSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYTGGTSIGALDIKTLLALSC HEQEMGYSSLVIGALL  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLRILRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRPDTKCLAA ATWGRLPGPETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLINITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF	ı							
KNKPPMELDLNSSSEDNKPGKRVRTNSRSTP TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKGLINBLNNLPVISNMTAALDSCSAA DGSLAAEMPKLEABGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFIDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLIRLIRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALPSGRGAATGGRQGGRPDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLERMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ	1	-		1	- 1			ARSAAAAPGSEASFTESRGLQNKNRGGANGK
TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKITDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKARADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLIRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALPSGRGAATGGRQGGRPDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ VNMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLERMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF		İ	ŀ					GRRGSLNASGRRTPPNCAAEDIKASPSSTNKR
YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKEK KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGFPDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ WMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFVGLFIPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ		l			1			
ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNVPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKERR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFILKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRKKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALPSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLFEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ							ľ	
SPGAGNPPGTPKGKRELMSNGPGSIIĞAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPPTPPQLIA PTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMI.GSLVLRKALAFRLLIRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ	l			- 1		•		
NSGKKKGLNNELNILPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTIDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFITTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKER KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLIRLIRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLIRMAQGCD FALGNDFLNITTKAQATKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ	ł		ł		ŀ			
DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFITTITGITPGLPSLTTITVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKEKEK KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRPDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ VMQEVSRNCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLIRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF	ı				- 1			
IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGFETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ WMAAALVHCYSKSPSNKDAALLEAARAQ WMAAALVHCYSKSPSNKDAALLEAARAQ TMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQATKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ		ŀ	l		ļ			
KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFIDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGFETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ VMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQATKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ					İ			
KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLIRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGFPDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ			ľ	ľ	l			
DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLIRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGFPDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ			1	l		.	Α.	
IKAEADKVYTFTDNAPSPSIGS  IKAEADKVYTFTDNAPSPSIGS  IKAEADKVYTFTDNAPSPSIGS  IKAEADKVYTFTDNAPSPSIGS  IKAEADKVYTFTDNAPSPSIGS  IKAEADKVYTFTDNAPSPSIGS  IKAEADKVYTFTDNAPSPSIGS  IKAEADKVYTFTDNAPSPSIGS  IKAEADKVYTFTDNAPSPSIGS  IKAEADKVYTFTDNAPSPSIGS  ITRGQPWRRAAAAGILPGREAAACLPSC/AS  VTAAVSGLLVGYELGIISGALLQIKTLLALSC  HEQEMGVSSLVIGALL  II40 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  IVAEQHGGASGRNVTTGSLGEPQWLRVATGG  RPGTSPALFSGRGAATGGRQGGRFDTKCLAA  ATWGRLPGPEETLPGQDSWNGVPSRAGLGMA  WPWAAALVHCYSKSPSNKDAALLEAARAQ  NMQEVSRNRCALLHSAAVQEYGYGN  II42 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD  FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC  CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF  GICKEYSRQ			]	ŀ	ļ	1		
1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGMA WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ								
VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRPDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ WNMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ	Г	139	2489	A	9234	207	443	
HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ WMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ	ĺ			. }		]		
1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLIRLIRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WPWAAALVVHCYSKSPSNKDAALLEAARAQ VMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ	L			}	I			HEQEMGVSSLVIGALL
1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRPDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ WMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ	_							MAQGNNYGQTSNGVADESPNMLVYRKV
RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ WMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ	ľ	141	2491	Α	9242	2	535	FVEAAVKMLGSLVLRRKALAPRLLLRLLRSP
ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WPWAAALVVHCYSKSPSNKDAALLEAARAQ\ NMQEVSRNRCALLHSAAVQEYGYGN  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD\ FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC\ CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF\ GICKEYSRQ				ſ		ĺ	[	
WPWAAALVVHCYSKSPSNKDAALLEAARAQ \(\text{NMQEVSRNRCALLHSAAVQEYGYGN}\)  1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD \(\text{FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC}\) \(\text{CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF}\) \(\text{GICKEYSRQ}\)	ĺ			ļ		İ	İ	
1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATILRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ			1	j		1		
1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ						]	1	
FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ	h	142	2492	A	9245	157	466	
CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ	ľ	-						
GICKEYSRQ								• -
	L						i	
	ſĪ	143	2493	A	9247	264	115	GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG
ARDSTSIIRMGPEIPPP	L						[	

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1144	2494	A	9260	1	401	KKVPGRLSEMSFSLNFTLPANTTSSPVTVDCGP SLGLAAGIPLLVATALLVALLFTLIHRRRSSIE AMEESDRPCEISEIDDNPKISENPRRSPTHEKN TMGAQEAHIYVKTVAGSEEPVHDRYRPTIEM ERRR
1145	2495	Α .	9264	175	411	METIWIYQFRLIEIGDSTVGKSCLLHRFTQGRF PGLRSPACDPTVGVDFFSRLLEIEPGKRIKLLL WDTAGQERFISIT
1146	2496	A	9277	592	814	MFTYLEGREGIKSQPKMEPHSVTRLECSGMI SAHCSLNLPGTSDSPASASR/VAGTTGMRHHA WLIFAFLVETGF
1147	2497	A	9279	1255	2	FRRGRRGEEKEEEEEEEGWVNGMENSHPP HHHHQQPPPQPGPSGERRNHHWRSYKLMIDP ALKKGHHKLYRYDGQHFSLAMSSNRPVEIVE DPRVVGIWTKNKE\LELSVPKFKIDEFYVDQV PPKQVTFAKLNDNIRENFLRDMCKKYGEVEE VEILYNPKTKKHLGIAKVVFATVRGAKDAVQ HLHSTSVMGNIIHVELDTKGETRMRFYEL\LV TGRYTPQTLPVGELDAVSPIVNETLQLSDALK RLKDGGLSAGCGSGSSSVTPNSGGTPFSQDTA YSSCRLDTPNSYG/QGTPLTPRLGTPFSQDSSY SSRQPTPSYLFSQDPAVTFKARRHESKFTDAY NRRHEHHYVHNSPAVTAVAGATAAFRGSSD LPFGTVGGTGGSSGPPFKAQPQDSATFAHTPP PAQATPAPGFR
1148	2498	A	9302	1026	6	IASIQNADTMPGVGLLVSHFSTLVSRQRCPNY ADPQNLTDVSIFLLLEVSGDPELQPVLAGLFL SMCLVTVLGNLLIILAISPDSHLHTPMYFFFSN LSLPDV\GFTSTTVPK\MIVDI\QSRSRVISYAG CLTQKSLFAIFGGTEE\NMLLSVMAYDRFVAI CHPLYHSAIMNPCFCAFLVLLSFFFLSLLDSQL HSWIVLQFTIIKNVEISNFVCDPSQLLKFACSD SIINSIFIYFHKDPERQLVLAGLFLSMCLVTVL GRLIIILDVSPDSHLPTPMYFFLSNLSLPDIGFT STTVPKMIVDIQSHGRVIFYAGCLTQMSLFAIF GGMEERHAPECDGL
1149	2499	A	9303	1	699	MASQEKDIFIGWGTIHLFRKPQRSFFGKLLRE FRLVAADRSMGRYMLFGVINLICTGFLLMWC SSTNSIALT\SYTYLTIFDLFSLMTCLISYWVTL RKPSPVYSFGFERLEVLAVFASTVLAQLGALF ILKESAERFLEQPEHTGRLLVGTFVALCFNLF TMLSIRNKPFAYVSEAASTSWLQEHVADLSR SLCGIIPGLSSIFLPRMNPFVLIDLAGAFALCIT YMLIEI
1150	2500	A.	9308	797	693	DRSTSVTRAGVQWCSLGSLQPRTPGLLRSSCL SLP
1151	2501	A	9309	205	406	VAIKELPVLWKWSKPTR\TAKEPPQTQQRAG SKTAAPPCQWSRMASEGPNIPCPGARHSDKQ FLICTI
1152	2502	A	9314	913	504	KPSPLITPPAVVLPPSAVLNLVNTFSSFPQVEV QGPLCGPRKGRLAVTIPFFGLS/LPKYMDHRR PPPHR\EIFFVFLAETGFHRASQAGPDLPTS/S/I PPTSA/FPKCWEYRSEPQCLPGCLSFSGILLDL GTNVSLRAA
1153	2503	A	9315	392	1	HPHRPRPGFRSPARSSRPCPVLTSLLPPFPSPSP PADDLVKAGRDRKDPQVR/ERRLRPNPGRLG GPR\PRPARARS/CHQPRLTRVCPRSPPPEARA PAPAAPARGRGAPKRNRPRTDTRAPRGSSAR PGNS

NO. of   No. of   N	SEQ ID	SEQ ID	Met	SEQ	Predicted	D-3:4-3-3	I Andrews and the second
Decidide   Seq						Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence			1				
1154   2504   A   9321   331   333   Methodine, N=Aspuragine, P=Proline, acid residue of residue of peptide sequence poptide sequence   Part Province, N=Aspuragine, P=A	eotide		}				l=Isolercine K=I vsine I=I encine
1160   2510   A   9341   1   390     1161   2511   A   9341   1   390     383     4   3941   1   390     390	seq-	uence	1	09/496	correspondi		M=Methionine, N=Asparagine, P=Proline
minino acid residue of peptide sequence   peptide sequence   peptide sequence   peptide sequence   peptide sequence   peptide   peptide sequence   peptide	uence		1	914			Q=Glutamine, R=Arginine, S=Serine.
	ł	ſ	Ì				T=Threonine, V=Valine, W=Tryptophan.
1154   2504   A   9321   331   333   MPCUAQYGTPAPSGPRDHSASDPLTPETEK   PT	1		ĺ	ĺ		sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1154   2504   A   9321   331   433			l	l			/=possible nucleotide deletion, \=possible
1155	1154	2504		0221		122	
1156	1154	2504	A	9321	331	433	MPCI/QAQYGTPAPSPGPRDHSASDPLTPEFIK
1156	1155	2505	A	9324	180	275	
1157   2507   A   9327   152   292   YERRORSQOGSSIPAROPORTOL ROPDIBLEPPESEP							MISPSRTEGIPI PLEDEGEGGERAVE
1157   2507   A   9327   152   292   YPERGESGGGGSEP_AGAOPGGRAIGAGWOS   KPPLWEGLORSGSP_AGGSE		]	j	]	]	1	EAAORHCRASVSII.RMRRPGOGSSRPARVPI
1157   2507   A   9327   152   292   YERRGRSQGGSHEAGAQPGGRAIGAGWQS   KEPLWBGLQRSGSPLPG							RGPDSHRLREPPPSPP
1158   2508   A   9328   1   430	1157	2507	Α	9327	152	292	YERRGRSQGGGSHPAGAOPGGRAIGAGWOS
LSKTFSVSSALAMLOERRCLTVVLITDSRCFL	I		L				KEPLWEGLQRSGSPLPG
VCMCFLTFIQALMVSGYLSSYITIERYSLKS	1158	2508	Α	9328	1	430	QELKQGPNPLAPSPSAPSTSAGLGDCNHRVD
SESGILVSCFDIGNLVVVVVSVSYRGRRRRP/ RVAAVGGILDLEGGEMS			l				LSKTFSVSSALAMLQERRCLYVVLTDSRCFL
1159   2509   A   9334   108   383   KGNOYNGNOLKRKHESMCPVSLTONTVE LMEAGLPQKQARRADELFRAGIUTYKLDER VLNALNYSVGLQWFKESDLSHLRILEISFT LMEAGLPQKQARRADELFRAGIUTYKLDER VLNALNYSVGLQWFKESDLSHLRILEISFT LAGFRYGERLRTAAM KRYVRILLIGEGAEHVADPYPGGRGVFRETAAM KRYVRILLIGEGAEHVADPYPGGRGVFRETAAM KRYVRILLIGEGAEHVADPYPGGRGVFRETAAM KRYVRILLIGEGAEHVADPYPGGRGVFRETAAM KRYVRILLIGEGAEHVADPYPGGRGVFRETAAM KRYVRILLIGEGAEHVADPYPGGRGVFRETAAM KRYVRILLIGEGAEHVADPYPGGRGVFRETAAM KRYVRILLIGEGAEHVADPYPGGRGWFRETAAM KRYVRILLIGEGAEHVADPYPGGRSW EAELPHMSQLTEIETCVEC CONTROL OF STANDAM AND C				l			VCMCFLTFIQALMVSGYLSSVITTIERRYSLKS
1169	ł			1			SESULLVSCFDIGNLVVVVFVSYFRGRRRRP/
1160   2510   A   9338   2   430	1159	2509	A -	9334	108	383	KGNOVNGNONON KDVHESMODVELTONITYD
1160   2510   A   9338   2   430						1000	LMEAGLPOKOAERADEL FEAGLVIVVKI DED
1160		<u>L</u>					VLNAL\YSSVGLOWFKESDLSHLRLLEISFR
KRYVRILLIGEGAEHVADPVPGGRGVPRGEA	1160	2510	Α	9338	2	430	FVGRPRGLSDRLEDLFLAGFRVGERLRTAAM
RTKWPLV/RWGDHA/SGPYCRISYLPSGRSM   EAELPIMSQLTEIETCVEC	[		}				KRYVRILLLGEGAEHVADPVPGGRGVPRGEA
1161   2511						•	DHTDQELREEIHKANVERVVHDVSOEATIEKI
1161   2511							RTKWIPLV/RWGDHA/EGPVGIKSYLPSGRSM
AVGICSPMSGVISMSAPFFLCKIIDATYTHYTV DYSDNLTRLCLGI.SGVFLCGAAANARVYLM QTSRQRVVKRLRTSLFSSILQQEVAFSDKAGT GELI  1162 2512 A 9343 84 837 QGRFRAFCWQRDFLQPFGMRLSALLALASKV TLPPHYRYGMSPPGSVADKRXNPPWIRRRPV VVEPISDEDWYLFCGDTVEILEGKDAGKQGK VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR GTMIPSEAPLLHRQVLVDPMDDREPTEIEWR FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET WIDGFKDTSVEDALERTYVPCLKTLQEEVME AMGIKETRNTRRSIGIEFQAEQLLPNFCPSLE G  1163 2513 A 9346 967 616 DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS LPSSWAYRCASPHPDNIFFVELVESGFHHVGQ AGLKLLISSDPPTSAFPKCWDYRRDSSAPAT FSSYQRNNPDLILNDTIMPNIK 1164 2514 A 9347 3 1099 SSPPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI HGGWRHHRDHTAIDEWDFNPSKELIYTCLLL FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGWWANPRYRTEGEACVEFKA MLIAVGHHLLLMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELILC SVNILQFIFIALKLDRIHWFWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVABQRRTH VTMAISWTITVPLLTFFVLLVHRLDGHNTFS YVSIFVPLWISLLTLMATTFRKGGCHHWWF AIRDFCQDQLPQPTGKPPPPELTDHHGEKA LPLQNKDRGSWPASRGSPRLL 1165 2515 A 9362 547 991 DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLAADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLTQKRC TKHAYGRRPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVETTSVHGRERLSAKD 1166 2516 A 9363 201 387 PPILRWIPPSGKNFFFFFFFFSEFFYSSPRVECS GAISAHLAHCHLCLPGSSDSPASAFQVAS	1161	2511	Α	03/1	,	200	
DYSDNI.TRLCLGLSGVFLCGAAANAIRVYILM QTSRQRVVKRLRTSLFSSILGQEVAFSDKAGT GELI  1162 2512 A 9343 84 837 QGFRAFCWQRDFLQPFGMRLSALLALASKV TLPPHYRYGMSPPGSVADKRKNPPWIRRRPV VVEPISDEDWYLFCGDTVEILLEGKDAGKQGK VVQVIRQRNWVVVGEINTHYRYIGKTMDYR GTMPSEAPLLHRQVKL.VDPMDRKPTEIEWR FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET WIDGPKDTSVEDALERTYVPCLKTLQEEVME AMGIKETRNTTRRSIGIEPGAEQLLPNFCPSLE G  1163 2513 A 9346 967 616 DSLALSPRLECSGAISAHCNLTTPGFTFFSCLS LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ AGLKLLISSDPPTSAFFKCWDYRRD\SSAPAT FSSYQRNNPDLILNDTIMPNIK 1164 2514 A 9347 3 1099 SSFPTCMRTVFHSNTSVSSLHRRGHVTPQLTI HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL FSVLLPIRLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MLIAVGIFLLLLMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFRSLDVVABQRRTH VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS YVSIFVPLWILSLITIMATTFRRKGGNHWWF AIRRDF/CQDQLPQFTGKPPPPTJDHHGEKA LPLQNKDRGSWPASRGSPRLL 1165 2515 A 9362 547 991 DVSIGPPLLRRPGSGREQTRSLSFPSDPESSFSP VPEGWILADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLTQKRC TKHAYGRKPIWLSQMACSGPEPTLHDCFFRP LGEDTLFHVEYTSVHGRRLSAKD PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS	1101	2211	Λ	7341	1	390	
1162 2512 A 9343 84 837 QGRFRAFCWQRDFLQPFGMRLSALLALASKV TLPPHYRYGMSPPOSVADKRKNPPWIRRPV VVEPISDEDWYLFGDTVEILEGKDAGKQGK VQVQVIRQRNWVVVGGLNTHYRYIGKTMDYR GTMPSEAPLLHRQVKLVDPMDRKPTEIEWR FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET WIDGPKDTSVEDALERTYVPCLKTLQEEVME AMGIKETRWTRRSIGIEPGAEQLLPNFCPSLE G  1163 2513 A 9346 967 616 DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS LPSSWAYRCASPHPDNFFVFLVESGFHHVQQ AGLKLLISSDPYTSAPFKCWDYRRDISSAPAT FSSYQRNNPDLILNDTIMPNIK  1164 2514 A 9347 3 1099 SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI HGGWRHHRDHTAIDEWDNPPSKFLYTCLLL FSVLLPLRLDGIIQWSYWAVPAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MLIAVGHLLLLMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSAACVWGFHDRSLELEILC SVNILQFIFIALKLDRIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFRSLDVVAAGRRTH VTMAISWITIVVPLLTFEVLLVIRLDGINTFS YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRRDF/CQDQLPQFTGKPPPFPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGFPLLRRPCSGREQTRSLSFFSDPESSFSP VPEGVRLADOPGHCRGRAVLTQKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFIVEYTSVHGRRRLSAKD PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS							
SELI   A   9343   84   837   QGRFRAFCWQRDFLQPFGMRLSALLALASKV   TLPFHYRYGMSPPGSVADKRKNPPWIRRRPV   VVBPISDEDWYLFCGDTVEILEGKDAGKQGK   VVQVIQRNWVVVGGLNTHYRYIGKTMDYR   GTMPSAFLHRQVKLVDPMDRKPTEIBWR   FTEAGERVRVSTRSGRIPPKPEFPPADGIVPET   WIDGPKDTSVEDALERTYVPCLKTLQEEVME   AMGIKETRINTRRSIGEIPGAEQLLPINFCPSLE   G   DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS   LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ   AGIKILISSDPPTSAFPKCWDYRRDSSAPAT   PSSYQRNIPPLILNDTIMPNIK   SSYQRNIPPLILNDTIMPNIK   SSYQRNIPPLILNDTIMPNIK   FSVLLPLRLJGIGWSYWAVFAPIWLWKLLV   VAGASVGAGVWARNPRYRTEGEACVEFKA   MLAVGIHLLILMFEVLVCDRVERGTHFWLL   FSVLLPLRLJGIGWSYWAVFAPIWLWKLLV   VVMAISWITLVPLLTFEVLLVHRLDGHNTFS   VVSIFVFLWLSLLTLMATTFRKGGNHWF   AIRRDF/CQDQLPQPTGKPPPPPTJDHHGEKA   LPLQNKDRGSWPASRGSPRLL   LPLQNKDRGSWPASRGSPRLL   LPLQNKDRGSWPASRGSPRL   TVMAISWITLVPLLTFEVLLVHRLDGHNTFS   VVSIFVFLWLSLLTLMATTFRKGGNHWWF   AIRRDF/CQDQLPQPTGKPPPPPTJDHHGEKA   LPLQNKDRGSWPASRGSPRL   TVMAISWITLVPLLTFEVLLVHRLDGHNTFS   VVSIFVFLWLSLLTLMATTFRKGGNHWWF   AIRRDF/CQDQLPQPTGKPPPPPTJDHHGEKA   LPLQNKDRGSWPASRGSPRL   TVMAISWITLVPLLTFEVLLVHRLDGHNTFS   VVSIFVFLWLSLLTLMATTFRKGGNHWWF   AIRRDF/CQDQLPQPTGKPPPPPTJDHHGEKA   LPLQNKDRGSWPASRGSPRL   TVMAISWITLVPLLTFEVLLVHRLDGHNTFS   VVSIFVFLWLSLLTLMATTFRKGGNHWWF   AIRRDF/CQDQLPQPTGKPPPPTJDHHGEKA   LPLQNKDRGSWPASRGSPRL   TVMAISWITLVPLLTFEVLLVHRLDGHNTFS   VVSIFVFLWLSLLTLMATTFRKGGNHWWF   AIRRDF/CQDQLPQPTGKPPPPTJDHHGEKA   LPLQNKDRGKWASGPEPTLHDCPFRP   LGEDTLFHVEYTSVHGRERJSAKD   PPILRWTPSGKNFFFFFFFSEFYSSPRVECS   GAISAHLAHCNLCLPGSSDSPASAFQVAS   GAISAHLAHCNLCLPGSSDSPASAFQVAS							OTSRORVVKRLRTSLESSILGOEVAFSDKAGT
TLPPHYRYGMSPPOSVADKRKNPPWIRRRPV VVEPISDEDWYLPCGDTVEILEGKDAGKQGK VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR GTMPSEAPLLHRQVKLVDPMDRKPTEIEBWR FTEAGERVRVSTRSGRIPKPEFPRADGIVPET WIDGPKDTSVEDALERTYVPCLKTI_GEVME AMGIKETRNTRRSIGIEFGAEQLLPNFCPSLE G  DSLALSPRLECSGAISAHCNLTPPGFTIPFSCLS LPSSWAYRCASPHEDNFFVFLVESGFHHVGQ AGLKLLISSDPPTSA/FPKCWDYRRD\SSAPAT FSSYQR\NPDLIL\nDTIMP\nik  1164 2514 A 9347 3 1099 SSPPTCMTYPHSNTSVSSLLHRPGHVTPQLTI HGGWRHRDHTAIDEWDF\PPSKFLIYTCLLL FSVLLPI.RLDGII\QWSYWAVFAPIWLWKLLV VAGASVGAGVWANPRYRTEGEACVEFKA MILAVGHILLLLMFEVL\VCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNII.QFIFIALKLDRIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVAGQRRTH VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRRDF\(\text{CQDQLPQPTQKPPPPFLTDHHGEKA}\) LPLQNKDRGSWPASRGSPRLL LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLRRPCSGREQTRSLS\PPSDFESSFSP VPEGVRLADGPGHCKGRVEVVKHQNQWYTV CQTGWSLRAAKVVCRQIRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPSKNFFFFFFESEFY\SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS			•				
VVEPISDEDWYLFCGDTVEILEGKDAGKQGK VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR GTMIPSEAPLLHRQVKLVDPMDRKPTEIBWR FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET WIDGPKDTSVEDALERTYYVFCLKTLQEEVME AMGIKETRINTRRSIGIEPGAEQLLPNFCPSLE G  1163 2513 A 9346 967 616 DSLALSPRIECSGAISAHCNLTPPGFTPFSCLS LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ AGLKLLISSDPTSA/FPKCWDYRRDNSSAPAT FSSYQRINPDLILINDTIMPNIK  1164 2514 A 9347 3 1099 SSPPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI HGGWRHRDHTAIDEWDFMPSKFLIVTCLLL FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MILIAVGHILLLIMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVAGQRRTH VTMAISWITIVVPLLTFEVLLVHRLDGHNTES YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGCKGRVEVVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLTQKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS	1162	2512	Α	9343	84	837	QGRFRAFCWQRDFLQPPGMRLSALLALASKV
VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR GTMIPSEAPILHRQVKLVDPMDRKPTEIEWR FTEAGERVRVSTRSGRIPKPEFPRADGIVPET WIDGPKDTSVEDALERTYVPCLKTLQEEVME AMGIKETRINTRRSIGIEPGAEQLLPNFCPSLE G DSLALSPRLECSGAISAHCNLTPPGFTFFSCLS LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ AGIKLLISSDPPTSA/FPKCWDYRRDSSAPAT FSSYGRNPPDLILINDTIMPNIK  1164 2514 A 9347 3 1099 SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV VGASSVGAGVWARNPRYRTEGEACVEFKA MLIAVGIHLLLMFEVLVCDRVERGTHFWILL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH VTMAISWITTVVPLLTTEFVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRRKGGNHWWF ARRDFCQDQLPQPTGKPPPPPTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRRFCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVILTQKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNIFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS							TLPPHYRYGMSPPGSVADKRKNPPWIRRRPV
GTMIPSEAPILHRQVKLVDPMDRKPTEIBWR FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET WIDGPKDTSVEDALERTYVPCLKTLQEEVME AMGIKETRINTRRSIGIEPGAEQLLPNFCPSLE G  1163 2513 A 9346 967 616 DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ AGLKLLISSDPPTSA/FPKCWDYRRDSSAPAT FSSYQRNPDLILINDTIMPNIK  1164 2514 A 9347 3 1099 SFPTCMRTVFHSNTSVSSLHRPGHVTPQLTI HGGWRHHRDHTADEWDFNPSKFLIYTCLLL FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MILIAVGHILLMFEVLVCDRVERGTHFWILL VFMPLFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH VIMAISWITIVVPLLTFEVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRKGGNHWWF AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPSGKNFFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASARQVAS					j		VVEPISDEDWYLFCGDTVEILEGKDAGKQGK
FTEAGERVRVSTRSGRIPKPEFPRADGIVPET WIDGPKDTSVEDALERTYVPCLKTLQEEVME AMGIKETR\NTRRSIGIEFGAEQLLPNFCPSLE G  1163 2513 A 9346 967 616 DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ AGLKLLISSDPPTSA/FPKCWDYRRDSSAPAT FSSYQRNNPDLILNDTIMPNIK  1164 2514 A 9347 3 1099 SSFPTCMTTVFISNTSVSSLLHRRGHVTPQLTI HGGWRHRDHTAIDEWDFNPSKFLIYTCLLL FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MLIAVGIHLLLIMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIIHWPWLVVVPPLWILM SFLCLVVLYTIVWSLLFRSLDVVAEQRRTH VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS YVSEVPI-WLSLLTLMATTFRKGGGNHWWF AIRRDF/CQDQLPQPTGKPPPPFLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRRPCSGRGTRSLSFPSDFSSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLTQKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS		100					VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR
WIDGPKDTSVEDALERTYVPCLKTLQEEVME AMGIKETR\NTRRSIGEPGAEQLLPNFCPSLE G  1163 2513 A 9346 967 616 DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ AGLKLLISSDPPTSAFPKCWDYRRDSSAPAT FSSYQRNNPDLILNDTIMPNIK  1164 2514 A 9347 3 1099 SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MILIVGHILLLMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLEIEILC SVNILQFIFIALKLDRIHWPWLVVFVPLWILM SFLCLVVLYIVWSLLFLRSLDVVABQRRTH VTMAISWTITVVPLLTFEVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD 1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS					ĺ		G1MIPSEAPLLHRQVKLVDPMDRKPIEIEWR
AMGIKETR\NTRRSIGIEPGAEQLLPNFCPSLE G  1163 2513 A 9346 967 616 DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ AGLKLLISSDPTSAFPKCWDYRRDSSAPAT FSSYQRNNPDLILNDTIMPNIK  1164 2514 A 9347 3 1099 SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MLIAVGHILLLLMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTILHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFFSSEFYSSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS							
C   DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS   LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ   AGLKILLISSDPPTSA/FPKCWDYRRDNSSAPAT   FSSYQRNPDLILNDTIMPNIK   SSPPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI   HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL   FSVLLPLRLDGHQWS YWAVFAPIWLWKLLV   VAGASVGAGVWARNPRYRTEGEACVEFKA   MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL   VFMPLFFVSPVSVAACVWGFRHDRSLELEILC   SVNILQFIFIALKLDRIIHWPULVVFVPLWILM   SFICLVVLYYIVWSLLFLRSLDVVAEQRRTH   VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS   YVSIFVPLWLSLLTLMATTFRKGGNHWWF   AIRRDF/CQDQLPQ/FTGKPPPPFLTDHHGEKA   LPLQNKDRGSWPASRGSPRLL		·					
LPSSWAYRCASPHPDNFFVELVESGFHHVGQ AGLKLLISSDPPTSA/FPKCWDYRRD\SSAPAT FSSYQRNNPDLILNDTIMPNIK  1164 2514 A 9347 3 1099 SSFPICMRTVFHSNTSVSSLLHRPGHVTPQLTI HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MLIAVGIHLLLIMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATIFRRKGGNHWWF ARRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFFSEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS					į		
AGLKLLISSDPPTSA/FPKCWDYRRD\SSAPAT FSSYQRNNPDLILNDTIMPNIK  1164 2514 A 9347 3 1099 SSPFTCMRTVFHSNTSVSSLLHRPGHVTPQLTI HGGWRHHRDHTAIDE WDFNPSKFLIYTCLLL FSVLLPLLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVAEQRTH VIMAISWITIVVPLLTFEVLLVHRLDGHNITFS YVSIFVPLWLSLLTILMATTFRRKGGNHWWF AIRRDF/CQDQLPQPTGKPPPPFLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS	1163	2513	Α .	9346	967	616	DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS
FSSYQRNNPDLILNDTIMPNIK  2514 A 9347 3 1099 SSFPICMRTVFHSNTSVSSLLHRPGHVTPQLTI HGGWRHHRDHTAIDEWDFNPSKFLIYICLLL FSVLLPIRLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILOFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVABQRRTH VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRKGGNHWWF AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLAADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLTVQKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFFSSEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS		}					LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ
1164 2514 A 9347 3 1099 SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTT HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLTVQKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS		1					
HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MLIAVGIHLLLMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVABQRRTH VTMAISWITTVVPLLTFEVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  DVSIGPPLLRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS	1164	2514		0247	3	1000	FSSYQRNNPDLILNDTIMPNIK
FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVABQRRTH VTMAISWITTVPLLTTEVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLTIQKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS		~17	^	7341	ا -	ללטו	SOFT I CMK I VEHSNISVSSLLHRPGHVTPQLTI
VAGASVGAGVWARNPRYRTEGEACVEFKA MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVABQRRTH VTMAISWITTVPLLTEFVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRKGGNHWWF AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLTQKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS							FSVI I DI DI DODOWOVWAVEADRUI WELLY
MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH VIMAISWITIVVPLLTFEVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLTVQKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD 1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS							VAGASVGAGVWARNDRVRTEGEACVERVA
VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH VTMAISWITTVVPLLTFEVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD 1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS		i					
SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVABQRRTH VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS	ſ	ľ		ľ	Ĭ		
VIMAISWITIVVPLLTFEVLLVHRLDGHNIFS YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS	ł		- 1				SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM
VIMAISWITIVVPLLTFEVLLVHRLDGHNIFS YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS	ļ		l				SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH
AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRGRAVLTIQKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS	ł	- 1	- 1	}	1		VIMAISWITTVVPLLTFEVLLVHRLDGHNTFS
LPLQNKDRGSWPASRGSPRLL  1165 2515 A 9362 547 991 DVSIGPPLLRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS	ļ	ļ		ŀ	l	. [	
1165 2515 A 9362 547 991 DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS	]						
VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD 1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS	1165	2515	$\overline{\mathbf{A}}$	9362	547	991	
CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS				7302		771	VPEGVRI ADGEGUCE CE VEUR LONGUERE
TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS	1	1	ĺ	- 1			
LGEDTLFHVEYTSVHGRERLSAKD  1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS	Ì				· I		
1166 2516 A 9363 201 387 PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS					ŀ	ł	
GAISAHLAHCNLCLPGSSDSPASAFQVAS	1166	2516	A	9363	201	387	PPILRWIPPSGKNFFFFFFESEFY/SSPRVECS
110/ 251/ A 9368 707 1087 AVLTPCLSPCSPSRIPRP\SRPYPGRRSLSHTPP	1167	2612					GAISAHLAHCNLCLPGSSDSPASAFQVAS
	110/	2017	A	9368	707	1087	AVLTPCLSPCSPSRIPRP\SRPYPGRRSLSHTPP

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamlc Acid,
nucl- eotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		÷-	914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine,
		ĺ	l	residue of	sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
1	ļ	}	l	peptide	buqaaas	/=possible nucleotide deletion, \=possible
L				sequence		nucleotide insertion
						PRPLILYAPAP\RPAGTAFIPHSHPPPPDLLRPT
İ		i		}		ATPA/TPCPSLPPPPRPLHPTQPSTALLPDPPPW
1168	2518	A	9375	511	16	PLPFPPPSS/RPPRPDCSTSYSPTFPPPT
1100	2010	^	93/3	311	15	MMLSEETSAVRPQKQTRFNGAKLVWMLKGS PITVTSAVIIVLMLLMM/IFSPWLATHDPNAID
		ĺ	ĺ			LTARLLPPSAAHWFGTDEVGRDLFSRVLVGS
		ł				QQSILAGLVVVATTGMIGSPLECLFGELGGRA
		}				DAIFMRYMDIMRS/IPSLVLTMEKTAALGPSL
						FNAMQASSEH
1169	2519	A	9377	42	410	GNGRVAPRDPGAVASAEPGLTTHDSGVNPN
	·					NSARRMEAMASGSNWLSGVNVVLVMAYWS LVFVLLFIFAKROIMRFAMKSLRGPHGPVGH
			·			NAPKOLKEEIDILLSRVHNIKYEPHILLADDDA
1170	2520	A	9378	302	1303	GVSGFSASVLRQRRMEDELEPSLRPRTQIQGR
						ILLLTICAAGIGGTFOFGYNLSIINAPTLHIOEF
]						TNETWQARTGEPLPDHLVLLMWSLIVSLYPL
						GGLFGALLAGPLAITLGRKKSLL\VNNIFVVS
						AAILFGFSRKAGSFEMIMLGRLASWGVNAGV SMNIQP\MLPGGESAPKELRGAVAMSSAIFTA
					•	LGIVMGQVVGLSTTAATGLRGL\AGELEELEE
						ERAACQGCRARRPWELFQHRALRRQVTSLV
						VLGSAMELCGNDSVYAYASSVFRKAGVPEA
}						KIQYAIIGTGSCELLTAVVSVSLEGALPPPAL
1171	2521	Α	9381	2	412	WGGTPRSFALNQFTLQKKKK
1 ****	2.721	^	3301	2	412	RGPASAQEDERARTAPLERVRARGRMTTSSA LFPSLLPCSWSTSNKYLAEFRAGKMSLKGTTE
						TPDKRKGLAY/IQQTDDSLIHFCWKDRTSGNV
	]			ļ		EDDLIIFPDDCEFKRLPQCPNGRVYVLKFKAG
1120	0500		0004			SKRLFFWMQEP
1172	2522	A	9384	20	355	GWNGRSTEASPAAEAPHVPHKETKAAMGTQ
	ľ			ŀ		CTHGGKVRPDPHDMLTTVVHKIKLFVLCHSL LQLCAIMISDYLKSSIYTVEKRLGLFRPTSGLL
						ASFNEVGNTALIVLESY
1173	2523	A	9393	430	87	LCQCIVPGQQKETFSLNPSSATVRFYL*LSLQ
	4					QRKEDQ*IIL*YHLNKDCLHIFMSAITLYMKI*
	- 1		}			KIPVLFDFNIMFETPFYII*FIFLFSQNLKRIRQV
1174	2524	A	9397	77	374	IRPPISFSKINNGP
****		•	7371	"	3/4	ERLEIGRLGGERGSGPASCLRVIDVSGMWDQ RLVKLALLQLLRAFYGIKVKGVRVHRDCGTF
	ł			ł	. 1	ESSSTLIRVS*FGVPCNALAHFGVTHF*YILDF
445						LGML
1175	2525	A	9399	66	397	HESSRADRDKMDTRGSTYTDADPVNKSGGT
	ĺ	i	1	ſ		AKMNKWSKGKVRDKLNNLVLFDTATYDKL
						CKEVPNYKLITLAVVSERLKIPGSLARAALHE LLSRGLI*LVIQHIAQVIY
1176	2526	A	9408	2	299	LDLTHVLSLSISLTVTILGTTFGMVIPLLDVVY
,			_ [	i		GERGYAQNGDF*DAQLDDYSFSCYSHAQVN
	i	ľ	i	1		GAPNSLTRAYDDP*VKISGLECQKVGALVEV
1125	2525	<u> </u>				KCLNL
1177	2527	A	9416	2	402	CNFLRSSRIRVHSTPAASTMPPKVDPNEIKVV
l		ł	1	ſ	·       [	YLRCTGGEVRATSALAPKIGPLGLSSIKVGVD
	ł	- 1	ł		1	FV*ATGDWNVLIISVILTIRILLSHIFVVPPFFCF DHLIAFWDLOSLIFLHVIFSLFITLLLFCFFSIF
1178	2528	A	9419	142	426	TPLFDLWPRVVLSWLETVLTSLRTRRAASGPP
l	ĺ			ł		ACRIMPTTVDDVLEHGGEVHFLQKQMLYLL
1130					1	ALI*DTFAPIYVGIVFLGFTPDHRCRSPGVAEL
1179	2529	A	9420	1450	1655	LSSAGTKMNLN*KNYWPGASAHACNPSTLG
1	1	ŀ		ļ	l	GQSRCITRSGDRDHPG*HGETPSVLKIQKISRA WWRAP
						M M LV/J

NO: of   NO: of   NO: of   node   Do NO:   node	SEQ ID	SEQ ID	1400	LCEO	Description	T 70 - 10 - 1	
Docation corticle			Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Decision   Sequence		b .	поа				D=Aspartic Acid, E=Glutamic Acid,
Sequence	1	1	{		1		
uence			l				
Inition acid   Inside of peptide sequence   T=Timeonine, V=Valine, W=Tryptopham, Y=Uprotine, X=Unknow, **-Skip codom, Y=Uprotine, X=Unknow, **-Skip codom, X=Deptide sequence   T=Timeonine, V=Valine, X=Unknow, **-Skip codom, X=Deptide sequence   T=Timeonine, V=Valine, X=Unknow, **-Skip codom, X=Deptide sequence   T=Timeonine, V=Valine, X=Unknow, **-Skip codom, X=Deptide sequence   T=Timeonine, V=Valine, X=Uproblem, X=Upro	-	uence					M=Methionine, N=Asparagine, P=Proline,
mino acid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence per per per per per per per per per pe	uence		l	914		acid residue	Q=Glutamine, R=Arginine, S=Serine,
residue of peptide sequence	1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryotophan,
Population   Pop	1	1		f	residue of	sequence	Y=Tyrosine, X=IJnknown, *=Stop codon.
1180   2530   A   9422   176   375   HRPQTTRFDWKPRT*PQGK*GRLSSEISPASPP SRFSRSTKFVPPKADPPARQKLTGVLHAPLLK L   1181   2531   A   9436   2   274   PHAASLRMYNLQPYTEENLICTAFATMVETVP HAPTILDRLTGPHGYCFVF*ADWATADKCVH TYNGKPLPAOTPILLS(L)AHLGS   1182   2532   A   9442   3   240   VDKCSSKSIVLSEYCPHCMCSLSTDPKFFGQL SMILK*MGAGDEKISAMGKARVDHRELYLGL LYPTEDYKLDTFARRH   1183   2533   A   9444   384   3   LKDFQFWALHDWFLFCCCTFLLFLVLECFTR KGCSGWAPWLSLQCQHGFRFRWADHLRSGV RDQFGQVSKTTFLFKIQKLAGHSGAHL*SYLL PERMR WKNRLNFGGRSCSEPRWHCHCTPGWAT ERG   RCMWWGAWFLSQCQHGFRFRWADHLRSGV RDQFGQVSKTTFLFKIQKLAGHSGAHL*SYLL REMR WKNRLNFGGRSCSEPRWHCHCTPGWAT ERG   LSGFKSLMFKPLQYTYVRVRTTWSFCLPLDG RKLMLS*YSK*LT*KYNLPEYSRMTLPPGMY INTCNSTLGGRAGWIV*AQEFST   LSGFKSLMFKPLQYTYVRVRTTWSFCLPLDG RKLMLS*YSK*LT*KYNLPEYSRMTLPPGMY INTCNSTLGGRAGWIV*AQEFST   LSGFKSLMFKPLQYTYVRVRTTWSFCLPLDG RKLMLS*YSK*LT*KYNLPEYSRMTLPPGMY INTCNSTLGGRAGWIV*AQEFST   RCPMWGQGASMDPAKNDRASTICCSLA WWWGWECWVRALKISSQPAGPLACWVAK KKSLSLSQPYPSBKGAGPCAGWACWLFGK KKSLSLSGPYPSBKGAGPCAGWACWLFGK KKSLSLSGPYPSBKGAGPCAGWACWLFGK KKSLSLSGPYPSBKGAGPCAGWACWLFGK KKSLSLSGPYPSBKGAGPCAGWACWLFGK KKSLSLSGPYPSBKGAGPCAGWACWLFGK KKSLSLSGPYPSBKGAGPCAGWACWLFGK KKSLSLSGPTPGTPQ WSAVVQFWLTAASNSCFSLLSSWDVYCA   HPQLHTKTHYVPTRMYNKI*QDDNSKFWQR GGTGLTHC**ESKLVQPLWKIVWHYQ   SSRGAPPCLSCNCHISPAFRKQRMGDSDQ*STT NPASPIPPEAPQEPWDSASGSVGSFSLGGAK ASS*VPGKGRGPQGSELLAETILELFLALAN SS*VPGKGRGPRQSELLAETILELFLALAN SS*VPGKGRGPRGPRADALSHMTGPMCLIENTT GRLMANPEALKLSATIQPMVEBALAGLYRAC **PYLTNNLAGMKKGLCLGSTEQAHIGIG RGVGSGPTAPLL PRQHCTLQTHRILHPEAPVKV*KT*PLFPGLR QASSCRRRCNPYLAARKAGSPRSHSTRENC RSSCCCHARASNLLYSSCQNHFGSPLAPLL PRQHCTLQTHRILHPEAPVKV*KT*PLFPGLR QASSCRRRCNPYLAARKAGSPRSHSTRENC RSSCSCLSVLRGGSSSNSHSFRRITTE MAAFVLLSVEQRPLKRPLQPPDVYPPDFQ KERELTAVNVK KUPPLFPDLR QASSCRRRCNFVLAARKAGSPSSHSTRENC RSSCSSNSSCSSNSHSFRRITTE MAAFVLLSVEQRPLKRPLQPPDVYPPDFQ KERELTAVNVK KUPPLFPDLR CREDERKMAREPLAEFMSTYVMMNHIMIVE KUPTSDLEHNTS		ì		1	peptide	•	
1180		ĺ	1	ſ		ĺ	nucleotide insertion
1181   2531   A   9436   2   274	1180	2530	Δ	0422		375	HDDOTTPDDWYDDT*DOCV*ODI CCCICD 4 CDD
1181   2531   A 9436   2   274	1100	2550	^	3422	170	3/3	
1181	i	1		i	i		
182   2532   A   9442   3   240   YDKCSSENSIVLSEYCPCMCSLSTDPKPFGQL   SMILK*MGAGDEKISAMGKARVDHRELYLGL   LYPIEDVKLTRRAH   1183   2533   A   9444   384   3   LKDFQFWALHDWFLFCCCTFLLFLVLECFTR   KGCSGWAPWLSLQCQHFGRPRWADHLRSGV   RQQFGQYSKTTFLPKIQKLAGHSGAHL*S*LL   ERMRWKNRLNPGGRSCSEPRWHHCTFGWAT   ERG   RCPKWQGQASKTFLFLKQKLAGHSGAHL*S*LL   ERMRWKNRLNPGGRSCSEPRWHHCTFGWAT   ERG   RCPKWQGQASRMDPAKAKDREASTICCSLA   RCPKWQGQASRMDPAKAKDRATICS   RCPKWQGQASRMDPAKAKDRATICS   RCPKWQGQASRMDPAKAKDRATICSLA   RCPKWQGQASRMDPAKAKDRATICS   RCPKWQGQASRMDPAKAKDRATICSLA   RCPKWQGQASRMDPAKAKDRATICSLA   RCPKWQGQASRMDPAKAKDRATICSLA   RCPKWQGQASRMDPAKAKDRATICSLA   RCPKWQGQASRMDPAKAKDRATICSLA   RCPKWQGQASRMDPAKAKDRATICSLA   RCPKWQGQASRMDPAKAKDRATICSLA   RCPKWQGQASRATICSL	1101	0531	<del>                                     </del>	0426 -	<u> </u>		
IYNGKPLPGATPLLSLQLHQLAHLGS	1 191	2531	A	9430	<sup>2</sup>	274	PIAASLRMYNLQPYTEENLICTAFATMVETVP
1182	1	l	!	1	ļ	1	
SMILK*MGAGDEKISAMGKARVDHRELYLGL   LYPIEDYKLITRARH				L			IYNGKPLPGATPLLSLQLHQLAHLGS
SMILK*MGAGDEKISAMGKARVDHRELYLĞIL LYPITEDYKLITRARH	1182	2532	A	9442	3	240	VDKCSSKSIVLSEYCPHCMCSLSTDPKPFGOL
LYPTEDYKLTFRARH		l	ĺ	!			SMILK*MGAGDEKISAMGKARVDHRELYLGL
1183		l					LYPTEDYKLTFRARH
KGCSGWAPWLSLQCQHFGRPRWADHLRSGV	1183	2533	A	9444	384	3	
RDQFQQYSKTTFLPKIQKLAGHSGAHL-95tL   ERMRWKNRLNPGGRSCSEPRWHCTPGWAT   ERG	i i					*	KCCCCM VDMI &I UCURECEDD MYDNI DCCA
ERMRWKNRLNPGGRSCSEPRWHHCTPGWAT   ERG							PDOPGOVENTEI PRIORI ACHICA ALII #C#1 1
ERG		1					
1184		•					
RKLMLS*YSK*LIT*KYNILPEYSRMTLPPGMV   HTCNPSTLGGRAGWIV*AQEFET	1104	2524		0460	201	655	
IHTCNPSTLGGRAGWIV*AQEFET	1104	2334	A	9462	391	655	
1185	[	[ .		1		· .	
WWWGWECWYRALKISSGFAGPLACWVAK   KKSLSLSGPVYPSEKGAGLYVF*DRVSLCHPG   WSAVVQFWLTAASNSCFSLLSSWDYRCA   WSAVVQFWLTAASNSCFSLLSSWDYRCA   HPQLITKTHYVPTRMVNKI*QDDNSKPWQR   GG*TGILTHCW*ESKLVQPLWKIVWHYQ   GG*TGILTHCW*ESKLVQPLWKIVWHYQ   SSRGAEPCLSNCHISPAPRKQRMGDSQP*SIP   NPASPHPEAPQEPWDSASGSVGSFSLGRGAK   ASS*VPGKGRGPRQGSELLAETILELFLALAN   S							
KKSLSLSGPYPSEKGAGLYVFDRVSLCHPG   WSAVVQFWLTAASNSCFSILLSSWDYRCA	1185	2535	Α	9467	215	566	RCPMWQGQASRMDPAKAKDREASTCCSLA
KKSLSLSGPYPSEKGAGLYVFDRVSLCHPG   WSAVVQFWLTAASNSCFSILLSSWDYRCA	1 1					·	WWWGWECWVRALKLSSGPAGPLACWVAK
WSAVVQFWLTAASNSCFSLLSSWDYRCA							
1186							
1187   2537   A   9469   388   3   EVAPGPSQILPRRVTDGGDRPQFSLPGPRLPQ   SSRGALPCLSNCIHSPAPRKQRMGDSDQ*STP   NPASPHPEAPQEPWDSASGSVGSFSLGRGAK   ASS*VPGKGRGPRQGSELLAETILELFLALAN   S     1188   2538   A   9471   124   397   TIMDKKNRHGNSLDMASEIHMTGPMCLIENTT   GRLMANPEALKILSAITQPMVEEAIAGLYRAC   *FYLTINNLAGMKKGLCLGSTEQAHTIGI   1189   2539   A   9480   584   769   GHVQSQHFGRPRRADHLRSGDRDHPG*HDET   PSLLKIQKISWAWWRAPVVPATWEAEAEEW   R     1190   2540   A   9483   463   86   VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL   PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR   GASSCRRRCNPYLAARKAGSPRSHSTRENC   RRSRCPDTAHRRRRGGRRNPSCVRSPRWR   1191   LADALCLSAAATGAVRPGARAQPSTRRILSP   SVRVCCRAAAASNLLYSSCLQRHSERASEEG   ERGSLSAKCSLVLRGGCSSSNSHSFRRIT*EI   MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ   KEERLTAVNVK   KEERLTAVNVK   CREDERKMAREFLAEFMSTYVMMNIHMIVE   KDTYSDHEEINTS	1186	2536	A	9468	275	452	
1187 2537 A 9469 388 3 EVAPGPSQILPRRVTDGGDRPQFSLPGPRLPQ SSRGAEPCL:SNCIHSPAPRKQRMGDSDQ*STP NPASPHPEAPQEPWDSASGSVGSFSLGRGAK ASS*VPGKGRGPRQGSELLAETILELFLALAN S  1188 2538 A 9471 124 397 TMDKKNRHGNSLDMASEIHMTGPMCLIENTT GRLMANPEALKILSAITQPMVEEAIAGLYRAC *FYLTINNLAGMKKGLCLGSTEQAHTIGI  1189 2539 A 9480 584 769 GHVQSQHFGRPRRADHLRSGDRDHPG*HDET PSLLKIQKISWAWWRAPVVPATWEAEAEEW R  1190 2540 A 9483 463 86 VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGRRRNPSCVRSPR WR  1191 2541 A 9489 1 411 LADALCLSAAATGAVRPGARAQPSTRRRLSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEBLTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CREDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS							
SSRGAEPCLSNCIHSPAPRKQRMGDSDQ*STP NPASPHPEAPQEPWDSASGSVGSFSLGRGAK ASS*VPGKGRGPRQGSELLAETILELFLALAN S  1188 2538 A 9471 124 397 TMDKKNRHGNSLDMASEIHMTGPMCLIENTT GRLMANPEALKILSAITQPMVEEAIAGLYRAC *FYLTNNLAGMKKGLCLGSTEQAHTIGI 1189 2539 A 9480 584 769 GHVQSQHFGRPRRADHLRSGDRDHPG*HDET PSILKIQKISWAWWRAPVVPATWEAEAEEW R  1190 2540 A 9483 463 86 VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGGRRNPSCVRSPRWR  1191 2541 A 9489 1 411 LADALCLSAAATGAVRPGARAQPSTRRRLSP SVRVCCRAAAASNILYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRTI*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEFELTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CREDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS	1187	2537	A	9469	388	3	
NPASPHPEAPQEPWDSASGSVGSFSLGRGAK ASS*VPGKGRGPRQGSELLAETILELFLALAN S  1188 2538 A 9471 124 397 TMDKKNRHGNSLDMASEIHMTGPMCLIENTT GRLMANPEALKILSAITQPMVEEAIAGLYRAC *FYLTNNLAGMKKGLCLGSTEQAHTIGI 1189 2539 A 9480 584 769 GHVQSQHFGRPRRADHLRSGDRDHPG*HDET PSLLKIQKISWAWWRAPVVPATWEAEAEEW R  1190 2540 A 9483 463 86 VTVGLTLLLRGAPRFTAG*PPSGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGRRRNPSCVRSPRWR  1191 2541 A 9489 1 LADALCLSAAATGAVRPGARAQPSTRRLSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEELTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS		1	**		500	,	
ASS*VPGKGRGPRQGSELLAETILELFLALAN S  1188 2538 A 9471 124 397 TMDKKNRHGNSLDMASEIHMTGPMCLIENTT GRLMANPEALKILSAITQPMVEEAIAGLYRAC *FYLTNNLAGMKKGLCLGSTEQAHTIGI  1189 2539 A 9480 584 769 GHVQSQHFGRPRRADHLRSGDRDHPG*HDET PSLLKIQKISWAWWRAPVVPATWEAEAEEW R  1190 2540 A 9483 463 86 VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGRRNPSCVRSPRWR  1191 2541 A 9489 1 411 LADALCLSAAATGAVRPGARAQPSTRRILSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEPELTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI*CHEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS							
1188 2538 A 9471 124 397 TMDKKNRHGNSLDMASEIHMTGPMCLIENTT GRLMANPEALKILSAITQPMVEEAIAGLYRAC *FYLTNNLAGMKKGLCLGSTEQAHTIGI 1189 2539 A 9480 584 769 GHVQSQHFGRPRADHLRSGDRDHPG*HDET PSLLKIQKISWAWWRAPVVPATWEAEAEEW R 1190 2540 A 9483 463 86 VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGRRNPSCVRSPRWR 1191 2541 A 9489 1 411 LADALCLSAAATGAVRPGARAQPSTRRLSP SVRVCCRAAAASNLLYSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEBLTAVNVK 1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI*CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS							
1188 2538 A 9471 124 397 TMDKKNRHGNSLDMASEIHMTGPMCLIENTT GRLMANPEALKILSAITQPMVEEAIAGLYRAC *FYLTNNLAGMKKGLCLGSTEQAHTIGI 1189 2539 A 9480 584 769 GHVQSQHFGRPRRADHLRSGDRDHPG*HDET PSLLKIQKISWAWWRAPVVPATWEAEAEEW R 1190 2540 A 9483 463 86 VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRRGRRNPSCVRSPRWR 1191 2541 A 9489 1 LADALCLSAAATGAVRPGARAQPSTRRLSP SVRVCCRAAAASNILLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEBLTAVNVK 1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CBEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS							ASS VPGAGRGPRQGSELLAETILELFLALAN
GRLMANPEALKILSAITOPMVEEAIAGLYRAC *FYLTNNLAGMKKGLCLGSTEQAHTIGI  1189 2539 A 9480 584 769 GHVQSQHFGRPRRADHLRSGDRDHPG*HDET PSLLKIQKISWAWWRAPVVPATWEAEAEEW R  1190 2540 A 9483 463 86 VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPYLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGGRRNPSCVRSPRWR  1191 2541 A 9489 1 LADALCLSAAATGAVRPGARAQPSTRRRLSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEBLTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS	1100	2520		0471	104	200	
1189 2539 A 9480 584 769 GHVQSQHFGRPRRADHLRSGDRDHPG*HDET PSLLKIQKISWAWWRAPVVPATWEAEAEEW R  1190 2540 A 9483 463 86 VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGRRNPSCVRSPRWR  1191 2541 A 9489 1 LADALCLSAAATGAVRPGARAQPSTRRLSP SVRVCCRAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEBLTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI*CHEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS	1100	2236	A	94/1	124	397	
1189 2539 A 9480 584 769 GHVQSQHFGRPRRADHLRSGDRDHPG*HDET PSLLKIQKISWAWWRAPVVPATWEAEAEEW R  1190 2540 A 9483 463 86 VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGRRNPSCVRSPRWR LADALCLSAAATGAVRPGARAQPSTRRLSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEBLTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI*CHEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS					-		GRLMANPEALKILSAITQPMVEEAIAGLYRAC
PSLLKIQKISWAWWRAPVVPATWEAEAEEW R  1190 2540 A 9483 463 86 VTVGLTLLLRGAPRFTAG*PPSGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGRRNPSCVRSPRWR  1191 2541 A 9489 1 LADALCLSAAATGAVRPGARAQPSTRRLSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEBLTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS							*FYLTNNLAGMKKGLCLGSTEQAHTIGI
1190 2540 A 9483 463 86 VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGGRRNPSCVRSPRWR  1191 2541 A 9489 1 LADALCLSAAATGAVRPGARAQPSTRRRLSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEBLTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI*CHEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS	1189	2539	A	9480	584	769	GHVQSQHFGRPRRADHLRSGDRDHPG*HDET
1190 2540 A 9483 463 86 VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGRRNPSCVRSPRWR  1191 2541 A 9489 1 LADALCLSAAATGAVRPGARAQPSTRRLSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEBLTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS							PSLLKIQKISWAWWRAPVVPATWEAEAEEW
PROHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRRGRRNPSCVRSPRWR  1191 2541 A 9489 1 LADALCLSAAATGAVRPGARAQPSTRRLSP SVRVCCRAAAASNILLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEBLTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS							R
PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGRRNPSCVRSPRWR  1191 2541 A 9489 1 LADALCLSAAATGAVRPGARAQPSTRRLSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEBLTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS	1190	2540	A	9483	463	86	VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL
GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGRRNPSCVRSPRWR  1191 2541 A 9489 1 LADALCLSAAATGAVRPGARAQPSTRRLSP SVRVCCRAAAASNILLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEBLTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS	1					1	
RRSRCPDTAHRRRRRGRRRNPSCVRSPRWR  1191 2541 A 9489 1 LADALCLSAAATGAVRPGARAQPSTRRLSP SVRVCCRAAAASNILLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEBLTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS	. 1						
1191 2541 A 9489 1 LADALCLSAAATGAVRPGARAQPSTRRRLSP SVRVCCRAAAASNILLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEELTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS	i .						
SVRVCCRAAASNILYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEELTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS	1191	2541	$\overline{\Delta}$	9489	1	A11	
ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEELTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS			••	7-07	- I	71.1	GADACCED V V V VOVIL I AGGCI CULTOCED VODDO
MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEELTAVNVK  1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS			• 1		l		
1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS	ļ		1		. I		
1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CHEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS	ì			1	·		
CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS	1100	0646		0.405			
KDTYSDHEEINTS	1192	2342	Α	9497	389	161	
KDTYSDHEEINTS			·		l		
1193   2543   A   9509   186   1   TAKSO*KRWORSGAMETI KHGUNDECT UCE							
I - I - I - I - I - I I I I I I I I I I	1193	2543	Α	9509	186	1	IAKSQ*KRWQRSGAMETLKHGWWECKLVQF
FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM		İ					FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM
1194 2544 A 9512 58 433 PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL	1194	2544	A	9512	58	433	
LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA					-		
SSKGQQFRR*KEHPPMLKTLNKLRIEGT*LKI		ľ	ı	İ			GGKCUUEDD & KDRDDDYU KALI YIAN DIDUARI ***
	ļ		[				
RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ  1195 2545 A 9515 595 1223 GHGAPSFOTOVPRTP*ASWPVVPAASESAPAP	1105	3545	<del></del> _	0515		1000	
I THE TAXABLE I SE	1133	2343	A	A212	טאַכ [	1223	GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP
AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC			- 1				AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC
PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP			- 1			_	PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP
LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL*	- 1		- 1				LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL*
PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP	ſ	. [	ſ	1		ł	PLHLLLHD*EKAWGFLFSSASHCFOGOICLIP
APGSGPCGATARPSRGGRAGGSRARRPIPPGP				ſ		1	APGSGPCGATARPSRGGRAGGSRARRPIPPGP
GTRTPSGCONPAASGG	I			- 1		. [	
						1	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Daniel and a	1 4 1 2 2 2
NO: of	NO: of	hod	ID NO:	beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	поп			nucleotide	D=Aspartic Acld, E=Glutamic Acid,
	,	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ł	l	914	ng to first	acid residue	Q=Ghitamine, R=Arginine, S=Serine,
1	l			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y≈Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
	İ			sequence	Į .	nucleotide insertion
1196	2546	A ·	9518	229	468	RSPTATPAPHAMGPGAPFARGGRPLPLLGAM
1					1	AERVAPGWDLHTPYLPRTNSRRTPHL**EPHA
	<u>L</u>	ĺ	(			GYIGALFPMSGGWPGGQ
1197	2547	A	9521	289	448	IAWLSGLFFPSNQANLCFLCYKLTADSRYRG
	ŀ	1			[	HAMRHLTGNTSMAIRFL*ADSRFQVQRARYE
1	i	İ			ĺ	APNWKYKYGY*IPVDMLC
1198	2548	A	9524	204	1	KNKKTTKCLSIVTLNISGPNQ*NKRHRVAEWI
i .	i	İ			Ī -	VKQEPNICHL*ETHFPFRDTYRLKEREQKKRK
1		1				SSYS
1199	2549	A	9546	1785	1943	GGRFKESKLTNAGWQRNSFFIGPPKSIPWAA
1			] , , , ,		1743	V*QRGDGKNPGVTHLNRPVGTX
1200	2550	A	9548	186	1	VNAEKEF*KIQHYFMTKSQNKLHIEHTYLKPI
		111	7540	100	* .	
1201	2551	A	9549	591	2	KAIYDKWTSDIMLNLQKL*AFFLRVIVRQI
1201	2,5,1	^ <u> </u>	3343	25,1	~	SSVVEFPRGPRSSLPPLDSTFPCGSSPNWTGGC
1			i			GSCPSGE*LVSPGSEQRKKYSNSNVIMHETSQ
						YHVQHLATFIMDKSEAITSVDDAIRKLVQLSS
[			i j			KEKIWTQEMLLQVNDQSLRLLDIESQEELEDF
						PLPTVQRSQTVLNQLRYPSVLLLVCQDSEQSK
j l						PDVHFFHCDEVEAELVHEYMESALTDCRLGK
1202	2652		0550	100		AMRP
1202	2552	A	9552	428	1	KYGNEGHWSRQCPNPGKPIRPCPLCRGPHWK
1						LDCERPPQGPLPSLPELAKTSYSDLTGLATED
}						*WGPGMDAPATTIASSKTRVTLMVAGRPVFF
1						LI*YRATYSALPNFSGPTQSSQVSVVGIDGQV
1000						SKPRATPPLFCSLHTF
1203	2553	A	9568	517	738	RRKFERKQKQ*RYREGKQYRQRDKMKEWG
1			1			EKEKRRREKGEREERKMRHRERKGESGQRD
100						TMENWRVERLTEKER
1204	2554	A	9573	83	415	EDKRLRLVDGDSRCAGRV*IYHDGFWGTICD
1 1						DGWDLSDAHVVCQKLGCGVAFNATVSAHFG
1		1		·		EGSGPIWLDDLNCTGTESHLWQCPSRGWGO
						HDCRHKEDAGVICSEFTALR
1205	2555	. <b>A</b>	9577	64	424	ARGSCPTRPRTANGRMGETKDAPQMLVTFK
						DVAVTFFREEWROLVLVHRTLYR*GMLETC
1					,	GLLDTLRHNVPQPDVVHLLYHGTQLLIVKRE
<u> </u>						VSHSPCAGDMRELFTREATLTPHPYNNGA
1206	2556	Α	9584	38	476	TLGAVLFSEVSKESSTSHSGGQLGRQNRHPKL
						SNFITPSSPRLKP*TASSQRNLGQILNMFLTAV
						NPQPLSTPSWQIETKYSTKVLTGNWMEERRK
[			ľ	1	ł	GLPYKHLITHHQEPPHRYLISTYDDHYNRHG
						YNPGLPPLRTWNGQKLLWL
1207	2557	A	9586	2	412	LRSSPAALLRALCITTVTGTALALRSRVATTN
				_	.1~	
ſĺ	- 1	. 1	Í			PDGCRNVLRPKYYRLCDKAESWGIALETVPT
		·		ļ		GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP
] ]	j	j	*			THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL
1208	2558	A	9597	122	<del>-</del>	FGILFSICFS
1400	٥٠٠	^	7371	144	3	IKNYWPGMVAHACNPSPLGGRGRWIA*AQK
1209	2550		06::	140		FADAWADAW
1407	2559	A	9611	148	558	KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK
	(	ſ	- (	i	i	GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ
	1	.	1	į		RIRDHDLLDKRKTVTALKAGEDRAILLGLAM
	)	J	J	1		MVCSIMM*FLLGITLLRSYMQSVWTRESQCT
197						LLNASITETFNC
1210	2560	A	9618	384	2	SLHDMLMLAEQQQKQKWAVNTQNTAWSNA
	ŀ			ľ		DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI
	ì		ł	1	i <b>l</b>	KVQVKNNDLGLQATINNEANWIAHQDDFNW
	ĺ	- 1	f	I	Í	LLAELNTCQRQETADS***WSPKNSHVGKDS
l					1	GELSAK
1211	2561	A	9620	316	610	QKHPGGGQLGRSPQEDSRFHNKASSGVSRVR
				<del></del>		YANG ANGRED TO ANGRED AND AND AND AND AND AND AND AND AND AN

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1.00	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	}	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	İ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1		/**	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1	Ì	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ŀ		İ	peptide	sequence	/=possible nucleotide deletion, \=possible
J	l	ł	) ·	sequence	1	nucleotide insertion
<b> </b>		<del> </del> -	<del> </del>	Sequence	<del> </del>	LGRAWWLTPVIPTLWEAKAGGSPE*D*AGRG
1					ľ	GSDI #SOUECDDDD MOUTE DE AMODODOOMOD
J		]		]		GSRL*SQHFGRPRRVDHLRSAVQDQPGQHGE TPSLLKIQKIN*VWGRRL*SSYSEAEAGESL
1212	2562	A	9623	297	344	QFPVDGDYQKIEKITQLFQAQNLSLCLAMTR
			7023	1	717	TREL*KGGGKGRHE*AVVPFLKKGGYGVKAP
j	ŀ	J				AILNTSNCT*CF*ETKMLSDDPKACVFEVSSA
	ł	ľ		j		DL*NTSFGVIR
1213	2563	A	9624	2	356	AELSLASTACGRNTSGDSLPDYDRAPISSPLA
i			, , , , ,	-	550	TSGTILSAISCLWDLPTPVLRVGLSCQPSMSSQ
l					1	PRMYSTDVEAAVNSLEDLYLQAYYAYLCVG
ŀ		ł				LYFHRDDMALEGVSRFL*ELAE
1214	2564	A	9634	776	912	SLSRWVRAKL*VPYNQENCLNPRGGGCSEPR
						SHYCTPAWATEKDS
1215	2565	A	9636	220	426	KPGNFAVSSEY*DITSGQLKTAVRG*IEMTST
					.20	EENFGEKLHDIGFGNGFLDKT*KAQATKAKI
		ł		İ		DK
1216	2566	A	9637	391	76	CFLEDGCTQAS*AEEAAVSPSMAEEEQGSTSC
		ł			1 ' '	RERRSIRFKMKNHSPDDTIKENVTISNIRTRKI
						NHLPETERNLLEHGLMYIRLNAAFCSLVAHS
	•					LFGFILKAT
1217	2567	Α	9655	2008	2432	LHCKMGALETQTHPCSQNMLRSLOKCCCKV
						EEHHLQPVQVLQTLLHSATAGTGCRRPARPP
						PAPPTPTPWRSRQSGKQSERAS*LKGRGRYGL
	]					GALGGRGGRALGGSRWPPPLPGETLFSGCKH
						RRRRGSDAAPGEEAGT
1218	2568	Α	9658	3	405	HASARALLSPNLSPNNKMAISGGPVLGFFIIA
						VLMSAQEPWAIKEEHVIIQAEFYLNPDQSGEF
						MLDFEGEDTFHGDMAKKETVWRLE*LARLD
						NFEAQRALANIAADQAALEIMDMGSDYTLIP
						NVPPKVTVL
1219	2569	Α	9662	3	284	PDWTEKRKMQDTGSILPLHWFGFGYAALVA
						YGGIIGYVKAGSVPSLAAGLLFGSLSGLGAYO
i					· ·	LSQDPRNVWVFLATSGTLAGIMGMRFYHSG
						_KL
1220	2570	Α	9669	200	699	LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSA
					1	PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN
		- 1	1	ļ	•	GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF
		ĺ				PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI
			ł	ł	1	SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL
122 I	2571		0676	777		K
1221	2711	A	9676	164	562	KERDSSTFSAAMTTMQGMEQAMPGAGPGVP
		-	1	· [	l	QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV
		i	ı			VQILTALMSLSMGITMMCMASNTYGSNPISV
. [		· [	ŀ	1	j	YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG
1222	2572	<del>_  </del>	9688	42	412	SLGMNITSS
	12 سا	A	2000	43	412	VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF
Í	1		ĺ	ľ		PTDENIKRKWVLAMKRLDVNAAGIWEPKKG
	}	İ	- 1	1		DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS
1223	2573	A	9696	308	564	PYHLQGKREKLHCRKNFTLKTVPATNYNH
	~,,	^	2020	200	564	RTSMGILYSEPICQAAYQNDFGQVWRWVKE
	j	J				DSSYANVQDGFNGDTPLICACRRGHVRIVSFL
1224	2574	A	9700	3	(22	LKKECLCQPQKPERENLLALCCE
	~/7	^	3100	•		DAWASGGELGSLFDHHVQRAVCDTRAKYRE
j	}	}	1			GRRPRAVKVYTINLESQYLLIQGVPAVGVMK
	1		•	[	[	ELVERFALYGAIEQYNALDEYPAEDFTEVYLI
						A PROBLEM SALE OF A CHARLES OF A CONTRACT A
.		1		}	1	KFMNLQSARTAKRKMDEQSFFGGLLHVCYA
.					1	PEFETVEETRKKLQMRKAYVVKTTENKDHY
						PEFETVEETRKKLQMRKAYVVKTTENKDHY VTKKKLVTEHKDTEDFRQDFHSEMSGFCKA ALNTSAGNSNPYLPYSCELPLCYFSSK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ĺ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
nence	[	ĺ	914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
		ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	Sequence	/=possible nucleotide deletion, \=possible
Į	!	1	j	sequence		nucleotide insertion
1225	2575	A	9710	1	163	RSGCVLRMTEWETGAPAVAETPDIKLFGKWS
Ĺ	L	L	_		1	TDDVHINDISLQDYIAGVRLILL
1226	2576	Α	9713	82	492	QGLPSFLPAFGPSGSWLGPAPTLGSSCNTVDT
	1		İ			ICHGYSEIRPLFYLSFCDLLLGLCWLTETLLYG
					ļ	ASVANKDIICYNLQAVGQIFYISSFLYTVNYI
	ļ				j	WYLYTELRMKHTQSGQSTSPLVIDYTCRVCQ
1227	2577		0700			MAFVFSSLI
1227	2311	Α	9720	3	416	GKWKRTQVPLLGEECADMDLARKEFLRGNG
						LAAGKMNISIDLDTNYAELVLNVGRVTLGEN
				·		NRKKMKDCQLRKQQNENVSRAVCALLNSGG GVIKAEVENKGYSYKKDGIGLDLENSFSNML
						PFVPNFLDFMQNGNYF
1228	2578	A	9723	278	411	EASSSNTVASNVADKTDPHSMNSRVFIGNLN
	,					TLVLQKSDVEAVF
1229	2579	Α	9725	121	902	LFAMSGFENLNTDFYQTSYSIDDQSQQSYDY
						GGSGGPYSKQYAGYDYSQQGRFVPPDMMQP
l i			1	1		QQPYTGQIYQPTQAYTPASPQPFYGNNFEDEP
					1	PLLEELGINFDHIWQKTLTVLHPLKVADGSIM
					•	NETDLAGPMVFCLAFGATLLLAGKIQFGYVY
						GISAIGCLGMFCLLNLMSMTGVSFGCVASVL
						GYCLLPMILLSSFAVIFSLQGMVGIILTAGIIG
						WCSFSASKIFISALAMEGQQLLVAYPCALLYG VFALISVF
1230	2580	Α	9739	11	247	TFVLNMNTPKEEFQDWPIVRIAAHLPDLIVYG
			2.02		- · ·	HFSPERPFMDYFDGVLMFVDISGKCKRDVCL
] ]				j		MWMSNRLAWEFTCRA
1231	2581	Α	9744	37	1100	TPLFDFWPGFVLSWLQPLSASLRARRAASGPP
				ĺ		ACRIMPTTVDDVLEHGGEFHFFQKQMFFLLA
						LLSATFAPIYVGIVFLGFTPDHRCRSPGVAELS
				•		LRCGWSPAEELNYTVPGPGPAGEASPRQCRR
						YEVDWNQSTFDCVDPLASLDTNRSRLPLGPC
				1		RDGWVYETPGSSIVTEFNLVCANSWMLDLFQ SSVNVGFFIGSMSIGYIADRFGRKLCLLTTVLI
						NAAAGVLMAISPTYTWMLIFRLIOGLVSKAG
						WLIGYILITEFVGRRYRRTVGIFYQVAYTVGL
		i				LVLAGVAYALPHWRWLQFTVALPNFFFLLY
				Į.		YWCIPESPRWLISQNKNAEAMRIIKHIAKKNG
						KSLPASL
1232	2582	Α	9753	164	517	PGPGMQGPPPITPTSWSLPPWRAYVAAAVLC
			į	1		YINLLNYMNWFIIAGVLLDIQEVFQISDNHAG
		1				LLQTVFVSCLLLSAPVFGYLGDRHSRKATMS
1233	2583	A	9757	25	419	FGILLWSGAGLSSSFISPRYSWLF LPAPWTERVRKSEGLVGTCLGDPMASPRTVT
		^	7131	<i>ــ</i>	717	IVALSVALGLFFVFMGTIKLTPRLSKDAYSEM
		I		1		KRAYKSYVRALPLLKKMGINSILLRKSIGALE
	J	· [	J	j	J	VACGIVMTLVPGRPKDVANFFLLLLVLAVLF
				1		FHOLV
1234	2584	A	9765	71	456	RLELDWGFSLHFLPVAYLCPLSSGFEMNVQP
	j	l		- 1		CSRCGYGVYPAEKISCIDQIWHKACFHCEVC
	ŀ	ì	Ì	1	l	KMMLSVNNFVSHQKKPYCHAHNPKNNTFTS
	ŀ	. [	ł	j		VYHTPLNLNVRTFPEAISGIHDQEDGEQCKSV
100 5						FHWD
1235	2585	A	9767	52	559	IRSGAMSVDKAELCGSLLTWLQTFHVPSPCA
1				i		SPQDLSSGLAVAYVLNQIDPSWFNEAWLQGI
	l	1		ļ	, ]	SEDPGPNWKLKVTSGLLIRGQTGEEMTRDGP
ł	ľ	- 1	ł	1	1	ARHMSWVMGRKRDRCLVINHLFIHSSMEYSP
			}			CARPGHSARNNTDKNLPHTAIILVTSNTYTTI
1236	2586	A	9770	352	608	KINFQAGRSGSCL FRGEAL TYPE TYPE IGD VASABLESTY FY IT C
			2.70	JJE	UV0	FRGEALTVRFLTKRFIGEYASNFESTYKKHLC

SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine.
seq-	uence	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ļ		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	<b>!</b>	l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			-	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l	i	peptide		/=possible nucleotide deletion_\=possible
		ĺ	1	sequence		nucleotide insertion
					i	LERKQLNLEIYDPCSQTQKAKFSLTSELHWA
						DGFVIVYDISDRSSFAFAKALI
1237	2587	Α	9793	266	515	NILAIIYFPFPRLFLLRDSQSNPKAFALTLCHH
		ŀ				QKIKNFQILPVSIDALTPPLVVCFLVSFLTHFS
					l	RYKPTRPVCITQFQGCS
1238	2588	Α	9802	537	967	ELGAGRSDREAMEAAVKEEISVEDEAVDKNI
	ŀ					FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT
	1	i				TDEDSTRFQUNEASKVPLLAEIYGIEGNIFRLK
	ļ	Ì		·		INEETPLKPRFEVPDVLTSKPSTVRLISCSGDT
						GSLILADGKGDLKC
1239	2589	Α	9805	105	540	VPGDPAMVRAGAVGAHLPASGLDIFGDLKK
		1				MNKRQLYYQVLNFAMIVSSALMIWKGLIVLT
	ŀ	1				GSESPIVVVLSGSMEPAFHRGDLLFLTNFRED
	· ·	l				PIRAGEIVVFKVEGRDIPIVHRVIKVHEKDNG
			·			DIKFLTKGDNNEGDDRGSYK
1240	2590	Α	9819	3	305	TDGRDPLPCAARRRGGGGECCGAGWVAEWS
						PQPLDPAMLLWMQGFVLEAVACQDNDDYLR
						YGILFEDLDCNGDGVVDIIELQEGLRNWSSAF
						DPNSEEHG
1241	2591	A	9834	841	1209	SPARGKSNRTDVMITAPKNKKMTENLAAPEA
						LDSSTHSSSTATQSRAKMNTPAPTPSTVPAIPR
						GGSGGPPPCAPHDRVSSVLQCDTQAMDHKTE
						SSHSVVEFLFKRTKTPSPFHPAVRENRN
1242	2592	A	9843	3	589	TISCGPATEPPASLLSSASSDDFCKEKTEDRYS
						LGSSLDSGMRTPLCRICFQGPEQGELLSPCRC
						DGSVKCTHQPCLIKWISERGCWSCELCYYKY
						HVIAISTKNPLQWQAISLTVIEKVQVAAAILGS
						LFLIASISWLIWSTFSPSARWQRQDLLFQICYG
				l		MYGFMDVMIVAVDSEDMVQAAKEVGKRWS
1010						DIPP
1243	2593	A	9846	198	411	WRISHHAGKMPVMKGLLAPQNIFLDTIATRF
						DGTHSNFILANAQVAKGFPIVYCSDGFCELAG
1011	0-04					PARTEVMQ
1244	2594	Α .	9848	116	650	PICGFLYLCSAMASESSPLLAYRLLGEEGVAL
				}		PANGAGGPGGASARKLSTFLGVVVPTVLSMF
						SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA
						LTVLSVCAIATNGAVQGGGAYCILQHRWTG
						VWPVLPAREVMISRTLGPEVGGSIGLMFYLA
1045	2606		0040		1200	NVCGCAVSLLGLVESVLDVFGA
1245	2595	A	9849	573	1620	KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE
						AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM
i				.	1	SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ
J						AMSLTEDVLAAALADHLPEDKWSAEKRRPL
1				ļ		KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR
			·	ļ		YVQPFLNALGAAGNFSVDSQILYYAMLGVNP
						RFDSASSSYYLDMHSLPHVINPVESRLGSSAA
	1	1		ł	ł	SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT
		- 1		ł		NAFHSPRWGGIMVYNVDSKTYNASVLPVRV
	ļ			l		EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL
	1	]	1	!	1	LSGPTSEGLMTWELDRLLWARSVENLATATT
1246	2505	<u></u> -l	0050	114	464	TLTSLA
1240	2596	A	9850	114	464	PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS
			1	l		ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM
			į		i i	QDGLSPCFFFTLVPSTRMALGTLALVLALPCK
						RRERPAGADSLSWGAGPRISSYV
1247	0.00			- I	327	FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF
1247	2597	A	9851	2	321	
1247	2597	A	9851	2	321	HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP
1247	2597	A	9851		321	

nucleotide contide control corresponding to the cor	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
DOSM   DOSM	NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
uence   91496   correspondi   ng to first   neith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith ceith residue   neith residue			İ				
Burne   Burn	seq-	uence	ļ-	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
	uence		İ	914			Q=Glutamine, R=Arginine, S=Serine,
Peptide	1		İ	ĺ			
1248   2598   A   9853   58			ł			sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1248   2598   A   9853   58							
1249   2599   A   9856   2   1265   LPPPEPERTRECEASTILATISGTONDOENE, VINLANDGILARSDLLINGDYTES VIGIBLITE LERIDEIITLLKNVGERVVLEVEYELPPPGGCP WT	1248	2598	A	9853		444	
1249   2599   A   9856   2   1265						```	
1249			İ				VSNLRPGGLAARSDLLNIGDYIRSVNGIHLTR
1259							LRHDEIITLLKNVGERVVLEVEYELPPPGGCP
RAPYRGQDSGAGTPQGRLAGRGAHLSRVGA   SGSUVAAGPAARHARRCADAGEAVOASC   GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT   PMGAGDAGASAESAVTTAPQEPPARRLQAGS   GAGPAGRAMRSTITLALLALLVLLYLVSGACL   PMGAGDAGASAESAVTTAPQEPPARRLQAGS   GAGPAGGRAMGSTITLALLALLVLLYLVSGAL   VFRALEQPHEQQAGRELGEVREKFLRAHPCV   SDQELGLLIEKVADALGGGGDVEHVFGSGKNEKFLRAHPCV   SDQELGLLIEKVADALGGGGDVEHVFGGKK   ELPHOGRCRETEGSQVAPRLPASPLCPGYGM   VALRTDAGRLFGIFYALVGIPLGGILAGVGD   RLGSSLRRIGIGHEAFLKWPVPELVRVLSA   MLFILIGGLLFVLTPTFVFCYMEDWSKLEATY   FVVYTLTTVGFGDYA   FVVYTLTTVGFGDYA   FVVYTDCPHOGRKELLAKVARLAMRDP   RLKWAVLVLVLVQMLACWLVRGLAWRWLL   FWAYAFGGCVNHSLTLAHDISHPAARFGCR   AARNRWLAVFANLPEGOFFVAASFKKVHVOH   HRYLGGDGLDVDVPTRLEGWPFCTARKLL   WLVQ.PFFYSLRPLCVJPFKAVASFKKVHVOH   HRYLGGDGLDVDVPTRLEGWPFCTARKLL   WLVQ.PFFYSLRPLCVJPFKAVASFKKVHVOH   HRYLGGDGLDVDVPTRLEGWPFCTARKLL   WLVQ.PFFYSLRPLCVJPFKAVASFKKVHVOH   HRYLGGDGLDVDVPTRLEGWPFCTARKLL   WLVQ.PFFYSLRPLCVJPFKAVASFKKVHVOH   HRYLGGDGLDVDVPTRLEGWPFCTARKLL   WLVQ.PFFYSLRPLCVJPFKAVASFKKVHVOH   HRYLGGDGLDVDVPTRLEGWPFCTARKLL   WLVQ.PFFYSLRPLCVJPFKAVASFKKVHVOH   HRYLGGDGLDVDVPTRLEGWPFCTARKLL   WLVQ.PFFYSLRPLCVJPFKAVASFKKVHVOH   HRYLGGDGLDVDVPTRLEGWPFCTARKLL   WLVQ.PFFYSLRPLCVJPFKAVASFKKVHVOH   HRYLGGDGLDVDVPTRLEGWPFCTARKLL   WLVQ.PFFYSLRPLCVJPFKAVASFKKVHVOH   HRYLGGDGLDVDVPTRLEGWPFCTARKLL   WLVQ.PFFYSLRPLCVJPFKAVASKKVHVOH   HRYLGGDGLDVDVPTRLEGWPFCTARKLL   WLVQ.PFFYSLRPLCVJPFKAVASKKVHVOH   HRYLGGDGLDVDVPTRLEGWPFCTARKLL   WLVQ.PFYSLRPLATKEVTRLTVTRLTVTRLTVTRLTVTRLTVTRLTVTRLTVTRL	1240	2500	<u> </u>	0956	2	1026	
SGSGVAAGPAARHAPRRCADAGEAVAGSC GRCAVALISGVCTLVSTHVCVGSGCPGAAGT GRCAVALISGVCTLVSTHVCVGSGCPGAAGT PMGAGDAGASAESAVTTAPQEPPARPLQAGS GAGPAGRAMARSTILLALLALLVILVSVGSALL VFRALEQPIEQQAQRELGEVREKFLRAHPCV SDQELGLIKEVADALGGGDWENGGK ELPHGGRCRETEGSQVAPRLPASPLCPGVGN VALRTDAGRIF-GEFYALVGIPLFGILLAGVGD RLGSSLRRGIGHEAIFLKWHVPPELVRUSA MLFLIGCLLFVLTFVFVCTWBDWSKLEATV FVIVILTTVFVFCWBDWSKLEATV FVIVILTTVTVFCWBDWSKLEATV FVIVILTTVGFGDYVA PRILAGVLEY STANDER SELVEN FRANKLIGHEAVALL STANDER SELVEN FVIVILTVOPHORANGASRTIMGNSASRNDF EWYYTDOPHORAKELAKVPAKALMRDP RLKWAV.VLVOMLAGWLOGLAWRWLL FWAYAFGGCVNHSLTLAHEDISPNAAFGTGR AARNRWLAVFANLPEGVPYAASFKKYNVOH HRVLGGDGLDVDVFTILEGWPFCTPARKLL WLVLQFFFYSLRFLCVIPFAVTROFTHRAKLL WLVLQFFFYSLRFLCVIPFAVTROFTHRAKL WLVLQFFFYSLRFLCVIPFAVTROFTHRAKL WLVLQFFFYSLRFLCVIPFAVTROFTHRAKL WLVLQFFFYSLRFLCVIPFAVTROFTHRAKL WLVLQFFYSLRFLCVIPFAVTROFTHRAKL WLVLQFFYSLRFLCVIPFAVTROFTHRAKL WLVLQFFYSLRFLCVIPFAVTROFTHRAKSK STANDERDEN SEVEN	1245	2399	^ <u> </u>	9630	2	1265	LPPPRPSKHRRGRAGTRASAAAAAGPTVSAV
GRCAVALISGVCTI.VSTHVCVGSGCGAAGT   PMGAGDAGASEANTTAPEPPARPLOAGS   GAGPAGRAMSTILIALIALVILYLYSGAL   VFRALEQPHEGOAGRELIGEVGCAGRELIGEVGERI.PHPCV   SDGE_GLI_LIKEVADAL_GGGADPETNSTNSSS   SAWDLGSAFFFSGTITTTIGGGGDWHVGGGK   ELPHGGRCRETEGSQVAPRIASELCTGYGN   VALRITDAGRICGTYALVGIFLGGILLAGVED   RUGSSLRRIGGHEAPILKWPVPEL.VSVLSA   MLPLLIGCLLFVLTPITFVFCYMEDWSKLEATY   FVVVTLTTVGFGDYVA   FVVVTLTTVGFGDYVA   FVVVTDCPTIVGFGDYVA   FVVVTDCPTIVGREVELAKYPAKALMRPDP   RLKWAVLVLVLVQMLACWLVRGLAWWLL   FWAYAFGGCVINSLITLAHDISHNAAFGTGR   AARNRWLAVPANLPEGVPYAASFKKYHVDH   HRYLGGGGLQVDVPTTRLEGWFCTPARKLL   WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV   QLA   ANRWILAVPANLPEGVPYAASFKKYHVDH   HRYLGGGGLQVDVPTTRLEGWFCTPARKLL   WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV   QLA   FVMPTLHESPGDIVEPSCCVSSSPELRRNAHSR   LBSYRPDTDLSREDTOCNLQHISDRINDDLN   MEPNPSDHPRASTIFLSKSGTDVREKRKSLFN   HHPPQQJARKYSSCSTIFLDDSTTVSQPNLKYTI   KCVALAYYHKNROPDGRMULDJEDENLHPL   SKSEVPPDYDKHPPEQKQYRFVRTLFSAAQL   TABCAIVTLVYLRELITLTAEDGANWKRIV   LGALLASKVWDDQAVWWDVQLLKDITVE   DMMELERQFLELLQFNINVPSATYAKYYPDL   RSLAEANNLSPLEPLSREHAHLLEAISRLCED   KYKDLRSARKRSASANDLTRENSPATIS   KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT   KAYGKMGMSLSTANPERCLUVYAKYTY   KAYGKMGMSLSTANPERCLUVYAKYTY   KAYGKMGMSLSTANPERCLUVYAKYTY   KAYGKMGMSLSTANPERCLUVYAKYTY   KAYGKMGMSLSTANPERCLUVYAKYTY   KAYGKMGMSLSTANPERCLUVAKYT   KAYGKMGMSLSTANPERCLUVAKYT   KAYGKMGMSLSTANPERCLUVAKYT   KAYGKMGMSLSTANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMS   KAYGKMGMSLATANPERCLUVAKYT   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGMS   KAYGKMGM	[ ]						SGSGVAAGPAARHAPRRRCADAGRAVGACC
PMGAGDAGASABSAVTTAPOEPPARLOAGS GAGRAPGRAMSSTILLALAULALVILVILVSGAL VFRALEOPHEQQAQRELGEVREKERABRCY SDQELGILLIKEVADALGGGADPETNSTSNSSH SAWDLGSAFFFSGTITTTIGGGDWHVGGGK ELPHGGRCRETEGSQVAPRLPASFLFOYGN VALRTDAGRIFGIFYALVGPILFGELLAGVGD RLGSSLRHGIGHIEAFILKWHVPFELVELSA ALLFLLIGGLLEVLTPTPTPCYMEDWSKLBAIY PVIVILTTVGFGDVVA  1250 2600 A 9873 2 652 PVVPSPCGGBFGRAPNGASSPTMGNSASRNDF EWYYTDQPHTQRRKEILAKYPAIKALMRPDF RLKWAVLVALVLQMLAVKJGLAWRWLL FWAYAFGGCVMHSLTLAHDISHNAAGTGR AARNEWLAVANLEGGDVALSKEILAKYPAIKALMRPDF RLKWAVLVALVLQMLAVSGLAWRWLL FWAYAFGGCVMHSLTLAHDISHNAAGTGR AARNEWLAVANLEGGDVALSKEILVDH HRYLGGDGLDVDVPTRLEGWFFCTPARKIL WLVLQPFFYSLRPLCVHFKAVTRIKVEILTH HRYLGGDGLDVDVPTRLEGWFFCTPARKIL WLVLQPFFYSLRPLCVHFKAVTRIKVEILTH HRYLGGDGLDVDVPTRLEGWFFCTPARKIL WLVLQPFFYSLRPLCVHFKAVTRIKVEILTH HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL WLVLQPFFYSLRPLCVHFKAVTRIKVEILTH HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL HRYLGGDGLDVDVPTRLEGWFFCTPARKIL KCVALATYHIKNRDPDGRMLDDFDNLHPT KCVALATYHIKNRDPDGRMLDDFDNLHPT LGABLASKASASADHVDQCDLKDTDVE DMNELRQFFLSCRSGSTANVEFTDVFKELSRICED LGABLASTANVERTUR LGABLASTAN LGABLASTANVERT							GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT
GAGRAPGRAMESTILALLALVILVISGAL VFRALEC/PIECOQARELEGVREKFLRAHPCV SDQBLGILLIKEVADALGGGADPETNSTSNSSH SAWDLGSAFFSGTITTIGGGGGOWHYVGGGK ELPHOGRCRETEGSQVAPRLPASPLCPGYGN VALRIDAGRIFCETYALVGIPLFGELLAGVGD RLGSSLRHGIGHEAFILKHPPFELVRVLSA MLFILIGCILFVALVGIPLFGELLAGVGD RLGSSLRHGIGHEAFILKHPPFELVRVLSA MLFILIGCILFVALVGIPLFGELLAGVGD RLGSSLRHGIGHEAFILKHPPFELVRVLSA MLFILIGCILFVALVGIPLAGVELAGVG RLGSSLRHGIGHEAFILKHPPFELVRVLSA MLFILIGCILFVALVGIPLAGVELAGVG RLGSSLRHGHEAFILGHPFELVSLSA MLFILIGCILFVALVGIPLAGVELAGVE RLGSSLRHGHPFELVSLSA MLFILIGCILFVALVGIPLAGVELAGVE RLX WAVLVLVIPLYGGDYALSPLACHAWRUL FWAYAFGGCVNHSLITLAHDISHNAARGTGR AARNRWLAVRANIPEGOPTAASFKKYHVDH HRYLGGDGLDVDVPTILLEGWFFCTFARKLL VILVLQPFPSLRPLCVHFRAVTIRMEVLINTLV QLA  1251 2601 A 9875 150 1209 PVIMPLHFSPGDIVRFSCCVSSSFKLRRNAHSR LBSYRPDTDLSREDTGCNLQHISDRERNDDLIN MEPHPSDHPRASTIFLSKSTDAVBRAVELTH KCVALAIYYHIKNRDPDGRMLDIDHRHRISLFH HHPPGQIARKYSSCSTIFLDDSTVSQPPLLYTH KCVALAIYYHIKNRDPDGRMLDIPDRIHIPT SKSEVPPDYDRINPERKYRSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPPLLYTH KCVALAIYYHIKNRDPDGRMLDIPDRIHIPT SKSEVPPDYDRINPERKYRSLFINK LGARLLAGNINLSPFLEPLSRERAHKLEAISRICED KYKDLRRSAARKRSAADNI-TI-PRWFAID  1252 2602 A 9879 6 376 KRPDSRPPAYRAGFTERGGELLYWKAT KAVGEMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSNVLLGARGETEEQLEKV WKSEVCSEPSSLSCSRSGKALISLSLYQ  1253 2603 A 9880 180 388 KEQAELLYGLYCCOCITI-SSHPSSVPAMSSC NFHAITVLIGIGLEKAHFWVGFPLLSMYVA AMFGNC  1254 2604 A 9881 19 494 VISFQITTDTIMDSSTAHSPVFLVPFPETTASEYE STELSATTFSTQSPLQKLFARKMKLIGTIQLP GIMTFSFQVIELFILLKYPRFPFETTASEYE STELSATTFSTQSPLQKLFARKMKLIGTIQLP GIMTFSFGVIELFILLKYPRFPFETTASEYE STELSATTFSTQSPLQKLFARKMKLIGTIQLP GIMTFSFGVIELFILLKYPRFPFETTASEYE STELSATTFSTQSPLQKLFARKMKLIGTIQLP GIMTFSFGVIELFILLKYPRFPFETTASEYE STELSATTFSTQSPLQKLFARKMKLIGTIQLP GIMTFSFGVIELFTLLKYPRFPFETTASEYE STELSATTFSTQSPLQKLFARKMKLIGTIQLP GIMTFSFGVIELFTLLKYPRFPFETTASEYE STELSATTFSTQSPLQKLFARKMKLIGTIQLP GIMTFSFGVIELFTLLKYPRFPFETTASEYE STELSATTFSTQSPLQKLFARKMKLIGTIQLP GIMTFSFGVIELFTLLKYPRFPFETTAGTT GATCVGLFPNGMCCPQAPFLGGGRGTDMMHPHPLT GATCVGLFPNGMCCPCAPFLGGGGRGTDMMHPHPLT GATCVGLFPNGMCCPCAPFLGGGGRGTDMMHP			1		•		PMGAGDAGASAESAVTTAPQEPPARPLOAGS
SDQELGLLIKEVADALGGGADPEINSINSISS	1 1						GAGPAPGRAMRSTTLLALLALVLLYLVSGAL
SAWDLGSAFFFSGTITTIGGGGGDWHYGGGK	i l						VFRALEQPHEQQAQRELGEVREKFLRAHPCV
						,	SDQELGLLIKEVADALGGGADPETNSTSNSSH
VALRTDAGRIECIPAL/GEIPEGILAGVGD							FI.PHGGRCRETEGGOVAPRI PASPI CPGVGNI
RIGSSLRHGIGHEAFILKWHVPPELVRVLSA   MFLIGGLLFVI-TKVFCYMEDWSKLEATY   FVIVILITVGGDYVA   MFLIGGLLFVI-TKVFCYMEDWSKLEATY   FVIVILITVGGDYVA   FVVFSPCGGIFGRAPNGASRFTMGNSASRNDF   EWYTTDQHTQRRKEILAKYPAIKALMRPDF   RILKWAVLVI-VLVQMLACWL-VRGLAWRWLL   FWAYAFGGCVNHSLAAFGTCGR   AARNRWLA VFANLFEGVPY-AASFKKYHVDH   HRYLGGDGLDVDVFTRLEGWFFCTPARKILL   WUVLQFFFYSLRFLCVHFKAVTRMEVLNTLV   QLA   QLA   Q875   150   1209   PVIMPLHFSPGDIVRFSCCVSSSPKLRRNAHSR   LESYRPDTDLSREDTGCALQHISDRENIDDLN   MEFNPSDHPRASTIFLSKQTDVREKRKSLFIN   HHPPQQIARKYSSCSTIFLDDSTVSQFNLKYTI   KCVALATYYHKNRDPDGRMILDIFDENLHPL   SKSEVPPDYDKHPPQKQNYRFVRTIFSAAQL   TAECAIVTLVYLERLLTVAEDICPANWKRIV   LGAILASKVWDDQ-VMVDVDYQOJLKDITVE   DMNELERQFLELLQFNINVPSSVYAKYYFDL   RSLAEANNLSFPLEPLSRERAHKLEAISRLCED   KYKDLRSARKRSADNLTLFRWSPAIIS   KRPDSRPPAQVRAGPTRPRTRGCELLYWKAT   KAVGIKMGSLSTANPECLDVFKELNSNNIG   DNIFFSSLSLLVALSMVLLGARGETEEQLEKV   WNSSEVCESPRSLSCSRGSSALLISLYQ   WNSSEVCESPRSLSCSRGSAKLILSLYQ   WNSSEVCESPRSLSCSRGSAKLILSLYQ   WNSSEVCESPRSLSCSRGSAKLILSLYQ   STELSATTFSTQSPLQKLFARKMKLLGTIQLIF   GMTTSFGVFFLYFFFFFFILSGYPFWG   SVIFTNSGAFLIAVKRKTTETLILSRIMNFLSA   LGAIAGILLTFEFFPRSKIHL   LGAILASKWDDGRGCTA   SPORT   SGGPAGLIAVVRKTTETLILSRIMNFLSA   LGAIAGILLTFEFFPRSKIHL   SQTPFWG   SVIFTNSGAFLIAVKRKTTETLILSRIMNFLSA   LGAIAGILLTFEFFPRSKIHL   LGATULASKVDDGRGCTTA   PHPIPAPEYTIQUT   PHPIPAPEYTIQUT   PHPIPAPEYTIQUT   PHPIPAPEYTIQUT   PHPIPAPEYTIQUT   PHPIPAPEYTIQUT   PHPIPAPEYTIQUT   PHPIPAPEYTIQUT   PHPIPAPEYTIQUT   PHPIPAPEYTICT   PHPIPAPEYTIGUT   PHPIPAPEYTI	Į						VALRTDAGRLFCIFYALVGIPLFGIT.LAGVGD
1250   2600   A   9873   2   652   FVVPSPCGGIPGRAPNGASRPIMGNSASRNDF	]						RLGSSLRHGIGHIEAIFLKWHVPPELVRVLSA
1250							
EWYTTO@HTQRRKELLAKYPAIKALMRPPP   RIKWAVLVLVLVQMLACWLVRGLAWRWLL   FWAYAFGGCVMTSLTLAIHDISHNAAFGTGR   AARNRWLAVFANLPEGPYAASFKKYHVDH   HRYLGGDGLDVDVPTRILEGWFFCTPARKLL   WLVLQPFYSLRPLCVIBKAVTRMEVLNTLV   QLA	1250	2600	Δ	0973		660	
RILKWAVLVLVLVQMLACWLVRGILAWRWLL   FWAYAFGGCVNHSLTLAHIDISHNAAFGTGR   AARNRWLAVFANLPEGVPYAASFKKYHVDH   HRYLGGDGLDVDVPTRLEGWFFCTPARKLL   WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV   QLA   QLA   QLA   Q875   150   1209   PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR   LBSYRPDTDLSREDTGCNLQHISDRENIDDLN   MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN   HHPPGQIARKYSSCSTITLDSTVSQPNLKYTI   KCVALATYYHIKNRDPDGRMLLDIFDENLHPL   SKSEVPPDYDKHNPEQKQUYRFVRTLFSAAQL   TAECAIVTLVYLERLLTYAEIDICFVANWKRIV   LGAILLASKVWDQAVWNVDYQILKDITVE   DMNELBEQFLEILQFNINVPSSVYAKYYFDL   RSLAEANNLSFPLEPJLSRERAHKLEAISRLCED   KYXDLRRSAKRSASADNLTLPRWSPAIIS   KRPDSRPPAQYRAGPTRPTRGCELLYWKAT   KAVGIKMGSLSTANVECLDVFKELNSNNIG   DNIFFSSLSLLYALSMVLLGARGETEEQLEKV   WNSSEVCSEPRSLSCSRSGSAKLILSLYQ   WNSSEVCSEPRSLSCSRSGSAKLILSLYQ   KRQLLYGGCDLTLTSSIPSSVPAMSSC   NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA   AMFGINC   AMFGINC   SULTINGAPILLAFTHILSRIMNFLSA   LGAIAGIILLTFEFHPRSKLHL   STELSATTFSTQSPLQKLFARKMKILGTIQILF   GIMTISFGVIFILTILKPYPRFFFFILSGYPFWG   SVLFINSGAFILLIS.RIMNFLSA   LGAIAGIILLTFEFHPRSKLHL   RGATCVGLPNVGKRTTETLILIS.RIMNFLSA   LGAIAGIILLTFEFHPRSKLHL   LGAIAGIILLTFEFHPRSKLHL   LGAIAGIILLTFEFHPRSKLHL   RATVAPDTMAQPYASAQFAPPQNGIPGEYTA   PHPHPAPEYTQGTT   GATCVGLPNVGMGLSGLLPILVPFILLG   GATCVGLPNVGMGLSGLLPILVPFILLG   GATCVGLPNVGMGLSGLLPILVPFILLG   GATCVGLPNVGMGLSGLLPILVPFILLG   GATCVGLPNVGMGLSGLLPILVPFILLG   GATCVGLPNVGMGLSGLLPILVPFILLG   GATCVGLPNVGMGLSGLLPILVPFILLG   GATCVGLPNVGMGLSGLLPILVPFILLG   GATCVGLPNVGMGLSGLLPILVPFILLG   GATCVGLPNVGMGLSGLLPILVPFILLG   GATCVGLPNVGMGLSGLLPILVPFILLG   GATCVGLPNVGMGLSGLLPILVPFILLG   GATCVGLPNVGMGLSGLLPILVPFILLG   GATCVGLPNVGMCCCPFSRGKKCLDFRKVSLTLYH   KEELE   EELE	1 1230	2000	^	3013	2	632	
FWAYAFGCWHISTLAHIDISINAAFGTGR	[						RLKWAVLVLVI.VOMLACWI.VRGI.AWRWI.I
AARRINALAVRANILEGUPYAASSKKYHVDH							FWAYAFGGCVNHSLTLAIHDISHNAAFGTGR
1251   2601   A   9875   150   1209   PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR   LESYRPDTDLSREDTGCNLQHISDRENIDDLN   MEFNPSDHPRASTIFLSKQTDVREKRKSLFIN   HHPPGQIARKYSSCSTITLDDSTVSQPNLKYTI   KCVALAIYYHIKNRDPDGRMILDIFDENLHPL   SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL   TAECAIVTLVYLERLLTYAEIDICPANWKRIV   LGAILLASKVWDDQAVWNVDYCQILKDITVE   DMNELERGLELLQFININVPSSVYAKYYFDL   RSLAEANNLSFFLEPLSRERAHKLEAISRLCED   KYKDLRRSARKRSASADNLTLPRWSPAIIS   KYRDLRSARKRSASADNLTLPRWSPAIIS   KYRDLRSARKRSASADNLTLPRWSPAIIS   KYRDLRSARKRSASADNLTLPRWSPAIIS   KYRDLRSARKRSASADNLTLPRWSPAIIS   KYRDLRSAR							AARNRWLAVFANLPEGVPYAASFKKYHVDH
1251   2601   A   9875   150   1209   PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR   LESYRPDITDLSREDITGCNLQHISDRENIDDLN   MEFNPSDHPRASTIFL,SKSQTDVREKRKSLFIN   HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI   KCVALAIYYHIKNRDPDEGMILDIPDENLHPIL   SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL   TAECAIVTLVYLERILTYAEIDLCPANWKRIV   LGAILLASKVWDDQAVWNDYQCQILKDITVE   DMNELERQFLELLQFNINVPSSVYAKYYFDL   RSLAEANNLSFPLEPLSRERAHKLEAISRLCED   KYKDLIRRSKKRSASADNITLPR WSPAIIS   CKYALIRRSKKRSASADNITLPR WSPAIIS   CKYALIRRSKKRSASADNITLPR WSPAIIS   CKYALIRRSKKRSASADNITLPR WSPAIIS   CHICAGO   DNIFFSISLIJALSMVLIQARGETEEQLEKV   WNSSEVCSEPRSLSCSRSSAKLILSLYQ   WNSSEVCSEPRSLSCSRSSAKLILSLYQ   WNSSEVCSEPRSLSCSRSSAKLILSLYQ   SWISSEVCSEPRSLSCSRSSAKLILSLYQ   A   9880   180   388   KEQABILYGLYCQCDLITJSHPSSVPAMSSC   NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA   AMFGNC   GIMTFSRGVIELFTILKPYPRFPFITASEYE   STELSATTFSTQSPLQKLFARKMKILGTIQILF   GIMTFSRGVIELFTILLKPYPRFPFITASEYE   STELSATTFSTQSPLQKLFARKMKILGTIQILF   GIMTFSRGVIELFTILLKPYPRFPFITILGGYPFWG   SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA   LGAIAGIILLTFEFHPRSKLHL   GATCVGLPNVGMCPQLSGALTFMYLQQGNQ   EATVAPDTMAQPYASAQFAPPQNGIPGEYTA   PHPHPAPETYGQTT   GATCVGLPNVGMCPQLSGALTFMYLQQGNQ   EATVAPDTMAQPYASAQFAPPQNGIPGEYTA   PHPHPAPETYGQTT   GATCVGLPNVGMCPQLSGALTFMYLQQGNQ   EATVAPDTMAQPYASAQFAPPQNGIPGEYTA   PHPHPAPETYGQTT   GATCVGLPNVGMCPQLSGALTFMYLQGNQ   EATVAPDTMAQPYASAQFAPPQNGIPGEYTA   PHPHPAPETYGQTT   GATCVGLPNVGMCPQLSGALTFMYLQGNQ   EATVAPDTMAQPYASAQFAPPQNGIPGEYTA   PHPHPAPETYGQTT   GATCVGLPNVGMCPQLSGALTFMYLQGNQ   EATVAPDTMAQPYASAQFAPPQNGIPGEYTA   PHPHPAPETYGQTT   GATCVGLPNVGMCPQLSGALTFMYLQGNQ   EATVAPDTMAQPYASAQFAPPQNGIPGEYTA   PHPHPAPETYGQTT   GATCVGLPNVGMCPQLSGALTFMYLQGNQ   EATVAPDTMAQPYASAQFAPPQNGIPGEYTA   PHPHPAPETYGQTT   GATCVGLPNVGMCPCCEVGEIDQCTK   PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH   KEELE   DQGPAPLAGGLGRPCPKKVSCTLYH   KEELE   DQGPAPLAGGLGRPCPKKVSCTLYH   KEELE   CACACACACACACACACACACACACACACACACACAC	1			•			HRYLGGDGLDVDVPTRLEGWFFCTPARKLL
1251   2601   A   9875   150   1209   PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR LESYAPDTDLSREDTGCNLQHISDRENIDDLN MEFNPSDHPRASTIFLSKSQTDVREKKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALATYYHIKNRDPDGRMILDIFDENLHPL SKSVEVPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLITYAEIDICPANWKRIV LGAILLASK WDDQA VWVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS KRPDSRPFAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ WNSSEVCSEPRSLSCSRSGSAKLILSLYQ WNSSEVCSEPRSLSCSRSGSAKLILSLYQ KEQAELLYGLYCQCDLTLSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC   VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQLF GIMTFSFGVIFLFILLKPYPRFFFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILTFEFHPRSKLHL GIGHT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAGQTYASAQPAPPQNGIPGEYTA PHPHPAPPYTGQTT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAGQTYASAQPAPPQNGIPGEYTA PHPHPAPPYTGQTT GIGHT GIGHT GATCVGLPNVGMCPQLSGALTFMYLQGGNQ DIGEPGHABGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE							
LESYRPDTDLSREDTGCNLQHISDRENIDDLN MEFNPSDHPRASTIFLSKSQTDVREKKSLFIN HHPPGQIARKYSSCSTIFLDDSTYSQPILKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPL SKSEVPPDYDKHNPEQKQIVIRFVRTLFSAAQL TAECAIVTLVYLERILLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTILFWSPAIIS KKPDSRPPAQVRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEJCDLVFKELNSNNIG DNIFFSSLSLLYALSMVLJGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ  1253 2603 A 9880 180 388 KEQAELLYGLYQCDLTLSSHPSSVPAMSSC NFIHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC  1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQLF GIMTFSFGVIFLFILLKPYPRFFFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREGROCTQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQPAPPQNGIPGEYTA PHPHPAPPYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHABGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE	1251	2601	A	9875	150	1209	
HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMILDIFDENLHPL SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFIELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS KYKDLRRSARKRSASADNLTLPRWSPAIIS KYKDLRRSARKRSASADNLTLPRWSPAIIS KYKDLRRSARKRSASADNLTLPRWSPAIIS KYKDLRRSARKRSASADNLTLPRWSPAIIS KYKDLRRSARKRSASADNLTLPRWSPAIIS KYKDLRRSARKRSASADNLTLPRWSPAIIS KYKDLRRSARKRSASADNLTLPRWSPAIIS KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ WNSSEVCSEPRSLSCSRSGSAKLILSLYQ KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFIHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC  1254 2604 A 9881 19 494 VISFQITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFILLKPYPRFFFFLSGYPFWG SVLFINSGAFLLAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIGPGHAEGILGKPCPKIKVSCEVEEIDQCTK PRDCEPNMKCCPFSRGKKCLDFRKVSLTLYH KEELE							LESYRPDTDLSREDTGCNLQHISDRENIDDLN
KCVALAIYYHIKNRPPGRMILDIFDENLHPL SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYXDLRRSARKRSASADNLTLPRWSPAIIS  1252 2602 A 9879 6 376 KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSINIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ  1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHAITYLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC  1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFILLKYPRFFFFFLSGYPFWG SVLFINSGAFLLAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVADTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQRPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE							MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN
SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTITPRWSPAIIS  1252 2602 A 9879 6 376 KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ  1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC  1254 2604 A 9881 19 494 VISFQITTDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIIL_TFEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQPAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHABGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE		- 0		1	ļ		HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI
TAECAIVTLYYLERLLTYAEIDICPANWKRIV LGAILLASKYWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS  1252 2602 A 9879 6 376 KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ  1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC  1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFFPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFFFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHFRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLFKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPFENMKCCPFSRGKKCLDFRKVSLTLYH KEELE				. }			KCVALAIYYHIKNKDPDGRMLLDIFDENLHPL
LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFI_ELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS  1252 2602 A 9879 6 376 KRPDSRPPAQYRAGPTRPRTRGCEILYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLIGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ  1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC  1254 2604 A 9881 19 494 VISFQITDTIMDSSTAHSPVFLVFPPEITASEYE GIMTTSFGVIFLFILLKPYPRFFFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRCCFQAPPLGLGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE							TARCAIVTI.VVI.ERI I TVAFIDICPANWEDIV
DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ  1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC  1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFILLKPYPRFFFIFLSGYPFWG SVLFINSGAFLIAVKRKTIETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE		J	]	•	J		LGAILLASKVWDDQAVWNVDYCOILKDITVE
RSLAEANNI.SFPLEPI.SRERAHKLEAISRI.CED KYKDI.RRSARKRSASADNI.TI.PRWSPAIIS  1252 2602 A 9879 6 376 KRPDSRPPAQYRAGPTRPRTRGCEILYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ  1253 2603 A 9880 180 388 KEQAELLYGLYQCODLTLSSIPSSVPAMSSC NFIHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC  1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE			ĺ				DMNELERQFLELLQFNINVPSSVYAKYYFDL
1252 2602 A 9879 6 376 KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEICLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCRSGSAKLILSLYQ  1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC  1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFFFIFLSGYPFWG SVLFINSGAFLIAVKRKTTEILIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE		i			Ī		RSLAEANNLSFPLEPLSRERAHKLEAISRLCED
KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ  1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC  1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATIFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTILKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE	1252	2602	<del></del>	0970	_	200	KYKDLRRSARKRSASADNLTLPRWSPAIIS
DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ  1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC  1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQLIF GIMTFSFGVIFLFILLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTKK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE		2002	^	2017	9	310	
WNSSEVCSEPRSLSCSRSGSAKLILSLYQ			ł				
1253   2603   A   9880   180   388   KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC     1254   2604   A   9881   19   494   VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFQVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL     1255   2605   A   9896   72   386   RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT     1256   2606   A   9902   95   399   SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE							
1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTILIKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTIETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE	1253	2603	Α	9880	180	388	KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC
1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATIFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFILLKPYPRFFFIFLSGYPFWG SVLFINSGAFLIAVKRKTIETLIILSRIMNFLSA LGAIAGIILTFEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE			- 1		Ì		NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA
STELSATTFSTQSPLQKLFARKMKILGTQILF GIMTFSFGVIFLFILLKPYPFFFFFILSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE	1254	2604		0001		404	
GIMTFSFGVIFLTILLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTIEITLIILSRIMNFLSA LGAIAGIIL.TFEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEBLE	12.54	2004	^	2001	13	474	VISTUITUTIMOSSTAHSPVFLVFPPEITASEYE
SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL  1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEBLE			}			\	
1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEBLE			}	l	l	j	
GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEBLE			ſ	[			LGAIAGIILLTFEFHPRSKLHL
GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE	1255	2605	A	9896	72	386	RPGREQRDCFQAPPLGLGGRQTDMMHHPLT
PHPHPAPEYTGQTT  1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFIILG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEBLE					ļ	1	GATCVGLPNVGMCPQLSGALTFMYLQQGNQ
1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEBLE	•	l	ł	- 1	- 1	1	
DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEBLE	1256	2606		99072	05	300	
PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEBLE	50		^	J,U2	,,	צענ	
KEBLE		j		i	Ī	i	PRDCPENMKCCPFSRGKKCLDFRKVSLTI VII
1257 10007 1 1 10005 1054							
1257   2607   A   9905   374   459   EHLKSTPNRLGVVAHTCNPSTLGGRGGW	1257	2607	Α	9905	374	459	EHLKSTPNRLGVVAHTCNPSTLGGRGGW

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	LA-i
NO: of	NO: of	hod	ID NO:	beginning		Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	100	in No.	nucleotide	nucleotide	D-Aspartic Acid, E-Glutamic Acid,
	1	]			location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=lsoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l	ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i	}	i	ľ	residue of	sequence.	Y=Tyrosine, X=Unknown, *=Stop codon,
	ì	1	l l	peptide	-	/=possible nucleotide deletion, \=possible
	1	ł	l	sequence	i	nucleotide insertion
1258	2608	A	9911	364	1974	
	2000	1	"""	304	1974	AGPGVPAVGGRWASGPGLGGRTLCSGPPDH
			Į.			QRRGPSCGASGDPQCVGSPHPQRARPLLARP
1	ł	1	ł	l	ł	GARLLPGHLPSPRPPRLPTGQPPAAAFRGPVR
1		l	ł		1	PQGGGHIHPLPTPGGRPCFAVSEGSGSALLLS
j	1	ļ	j	1	ł	YLGECGSSSYVTGAACISPVLRCREWFEAGLP
i		1	ļ		1	WPYERGFLLHQKIALSRYATALEDTVDTSRL
1		ĺ		1		FRSRSLREFEEALFCHTKSFPISWDAYWDRND
1	1	Î	1	[	1	PLRDVDEAAVPVLCICSADDPVCGPPDHTLTT
ì	1	i	i	i		ELFHSNPYFFLLLSRHGGHCGFLRQEPLPAWS
1	ľ	l	l	ł	ł	UEVII PERDAI TEREPTERRIKOI TEREPLANS
	l	Į.		ļ		HEVILESFRALTEFFRTEERIKGLSRHRASFLG
1	ł	ł	į	ì	1	GRRRGGALQRREVSSSSNLEEIFNWKRSYTRL
	İ	l	1		l	MAAAAGAAAAPGSREPQDRPECGAGHPGPR
-	-	1		ļ	ŀ	YYRHPERWLLRPEAFLGPLRTRAPSAEDSQR
		ŀ			j	ERPAARSGPEMRVRYPVVAAVLAPYLALSQD
l l	ł	ł	}		1	PMVKSSASGQGASGSYNHVREEMLIKAGGA
ľ				ĺ	i '	MSRRVVRQSKFRHVFGQAAKADQAYEDIRV
1		}	J		,	SKVTWDSSFCAVNPKFLAIIVEAGGGGAFIVL
			1	-		PLAK
1259	2609	A	9919	693	935	
	200)	1	1 //1/	095		GCFKFIGESTCCWIFPSSVTTQCVVAKAPRAA
1					· ·	TLSKAERLRSQPGPEQGGSSYRPRTPTAAAIL
1260	2610	_	2021	455	1000	PPRPGRSHRKRKLVSTK
1200	2610	A	9921	455	1082	QRSCLCSAIEKDGGDVKALYRRSQALEKLGR
						LDQAVLDLQRCVSLEPKNKVFQEALRNIGGQ
1 1						IQEKVRYMSSTDAKVEQMFQILLDPEEKGTE
1 1						KKQKASQNLVVLAREDAGAEKIFRSNGVQLL
			i l			QRLLDMGETDLMLAALRTLVGICSEHQSRTV
1 1						ATLSILGTRRVVSILGVESQAVSLAACHLLQV
1 1						MFDALKEGVKKGFRGKEGAIIV
1261	2611	A	9928	1	438	GFRGAEAPGAAQAPKKKKPRPTEGGPGAGSG
! · i			,,,,,	•	750	DCVDDVDCDTT I HODVDDVDCDT OU DCVDV D
j j				l		RGKDPYRGPTLLHQPKPPKDEFLSSLESYEIAF
1 1			[			PTRVDHNGALLAFSPPPPQRQRRGTGATAES
1 [						RLFYKEASPSTHFLLNLTRSSRLLAGHVSVEY
1262	2610					WTREGLAWQRADRPHCLYA
1262	2612	A	9931	168	435	AAEMGRAGAAAVIPGLALLWAVGLGGPPPA
1 1	- 1	- 1		ł	1	PPRLPFCLQELQGRHALHTFSLERTCSYQDFL
		í				WADEGRLLHVGAQDLATWHTLSPLGLW
1263	2613	A	9938	247	488	RMSATSVDQRPKGQGNKVSVQNGSIHQKDG
l I				100	,,,,	CNDDDFEPYLRSPDNQSNSYPPMSDPYMPGY
ļ 1	ł	ŀ		1	}	
1264	2614	A	9941	61	2777	YAPSIGFPYSLGEAAWSQL
			//TL	<b>71</b>	277	ESIGLTALGPRRRPWEHRWSDPITLKMKGWG
				ļ		WLALLLGALLGTAWARRSQDLHCGACKAVR
1265	<u> </u>	<u>, —</u> І				RRVRQFNIYDY
1265	2615	A [	9956	2	522	FVASEVSKMPVPASWPHPPGPFLLLTLLLGLT
	İ	l	- 1	l		EVAGEEELQMIQPEKLLLVTVGKTATLHCTV
ľ	Ì		1		1	TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP
j ł	ì	·			ł	RVTTVSDLTKRNNMDFSIRISSITPADVGTYY
<i>i</i>		ļ	1	ŀ		CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG
	1	i	. [			FLSOVWWWLSSHPFMN
1266	2616	A	10002	242	302	
.200	2010	^	10002	243	387	PKNNACHLLFTAVCQPRCKHGECIGPNKCKC
1000	0615	1				HPGYAGKTCNQGRKTV
1267	2617	A	10004	36	707	LPAPASTWSVARETMASSSVPPATVSAATAG
1	}	J	1	1		PGPGFGFASKTKKKHFVQQKVKVFRAADPLV
İ	ĺ		j		j	GVFLWGVAHSINELSOVPPPVMLLPDDFKAS
l		i	1		. I	SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL
	{	ſ	- 1	ſ	ļ	RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS
1		ŀ	1			ADDIT AIRERGEDIADI AND AND AND AND AND AND AND AND AND AND
1		ł	i	J		YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS
1268	2610	<del></del> +	10005			SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP
1500	2618	A	10005	2	209	GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP
				1		SQDELEHSLGESAAQGAAGVVLWVSWENTR

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
	!			amino acid residue of peptide	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				sequence		/=possible nucleotide deletion, \=possible nucleotide insertion TKVSLGLA
1269	2619	A	10010	245	688	FGMLKNKGHSSKKDNLAVNAVALQDHILHD LQLRNLSVADHSKTQVQKKENKSLKRDTKAI IDTGLKKTTQCPKLEDSEKEYVLDPKPPPLTL AQKLGLIGPPPPPLSSDEWEKVKQRSLLQGDS
1270	2620	A	10011	2.	588	VQPCPICKEEFELRPQVFSIRG RVDDFVRPLPPGLMSRSRASIHRGSIPAMSYA
			·			PFRDVRGPSTHRTQYVHSPYDRPGWNPRFCII SGNQLLMLDEDEIHPLLIRDRRSESSRNKLLR RTVSVPVEGRPHGEHEYHLGRSRRKSVPGGK QYSMEGAPAAPFRPSQGFLSRRLKSSIKRTKS
						QPKLDRTSSFRQILPRFRSADHDRYRGWSMW DEIDV
1271	2621	A	10013	209	363	LPAPPNLSPRLSFGFQFPGGNDNYLTTTGPSHP FLSGAEVSQSCRRRGGRA
12/2	2022	A	10014	7	388	SAVTISWKWRSVMGIQTSPALLASLGAGLVT LLGLAVGSYLVRRSRRPQVTLLDPNEKDLLR LIDKTLSARSPCKHIYLSTRIDGSLSIRPYTPVT SDEDQGYVDIDIKVYLKGVHPTFPEGGKMSH
1273	2623	A	10016	1	1339	MAARTLGRGVGRLLGSLRGLSGQPARPPCGV SAPRRAASGPSGSAPAVAAAAAQPGSYPALS AQAAREPAAFWGPLARDTLVWDTPYHTVW
				1		DCDFSTGKIGWFLGGQLNVSVNCLDQHVRKS PESVALIWERDEPGTEVRITYRELLETTCRLA NTLKRHGVHRGDRVAIYMPVSPLAVAAMLA
						CARIGAVHTVIFAGFSAESLAGRINDAKCKVV ITFNQGLRGGRVVELKKIVDEAVKHCPTVQH VLVAHRTDNKVHMGDLDVPLEQEMAKEDP
·						VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITGHSYVVYGPLCNGATSVLFESTPVYPNA GRYWETVERLKINQFYGAPTAVRLLLKYGD
			_			AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT
1274	2624	A	10017	1	3750	FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE
					0	KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV
		ļ				VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGIKDTVTTGVMGAVNLAKGTVQTGVETS
		. 1				KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT
			-			IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG
						AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD
						AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA
					1	KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTIYSGVTSAVNVAKGAVOT
					1	GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK GTVQTSVDTTKTVLTGTKDTVCSGVTGAAN

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F-Phenylalanine, G-Glycine, H-Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ľ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	i		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	l		ŀ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		I	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	Ĭ	ĺ	ſ	peptide	Jongonso	/=possible nucleotide deletion. \=possible
}	1	!		sequence		nucleotide insertion
	<del>                                     </del>			Boquano	<del>}</del>	VAKGAIQGGLDTTKSVLTGTKDAVSTGLTGA
1	j	1	ļ	1	ı	VKLAKGTVQTGMDTTKTVLTGTKDAVCSGV
			1	·		TGAANVAKGAVQMGVDTAKTVLTGTKDTV
1		j	1			
ł	1	ļ	}		ł	CSGVTGAANVAKGAVQTGLKTTQNIATGTK NTLGSGVTGAAKVAKGAVQGGLDTTKSVLT
1	ł	ļ	ł		į.	GTKDAVSTGLTGAVNLAKGTVQTGVDTSKT
	1		Ī	'	}	VLTGTKDTVCSGVTGAVNVAKGTVQTGVDT
İ	ì	i		!		AKTVLSGAKDAVTTGVTGAVNVAKGTVOTG
[	1	1	1	1	1	VDASKAVLMGTKDTVFSGVTGAMSMAKGA
ĺ	i	(,	ļ	1		VQGGLDTTKTVLTGTKDAVSAGLMGSGNVA
1		i '	]	1	1	TGATHTGLSTFQNWLPSTPATSWGGLTSSRT
ļ			]	1	J	TDNGGEQTALSPOEAPFSGISTPPDVLSVGPEP
I	1	i '		, ,		AWEAAATTKGLATDVATFTQGAAPGREDTG
]			1	1 ' !	ſ	LLATTHGPEEAPRLAMLQNELEGLGDIFHPM
1		!	!			NAEEQAQLAASQPGPKVLSAEQGSYFVRLGD
1				[ ]	1	LGPSFRQRAFEHAVSHLQHGQFQARDTLAQL
1						ODCFRL ODCFRL
1275	2625	A	10025	124	415	TILARKKEKTCPCKKEIGRNSRSGMYSRKAM
						YKRKYSAANTKVEKKKKEKVLAPVTKPVGG
1	İ			1		DKNGGTRVVKLPTMPRYYPTEDVPRKLLSHG
l	l i			1		KKPFS
1276	2626	A	10030	3	507	GGSLRFSPPRVPSCSRVFCPVPPGGCGLPSPMS
	-52-5		10050	1	30,	ASRPQSPTTPWCLPRRYMKHKRDDGPEKQED
						EAVDVTPVMTCVFVVMCCSMLVLLYYFYDL
1	[ !		ĺ	1 1		LVYVVIGIFCLASATGLYSCLAPCVRRLPFGK
ĺ				i !		CRIPNNSLPYFHKRPQARMLLLALFCVAVSV
						VWGVFRNEDQ
1277	2627	Α	10035	51	869	YSRFTVPLPATMASSEVARHLLFQSHMATKT
	]	J		j	]	TCMSSQGSDDEQIKRENIRSLTMSGHVGFESL
	<u> </u>	1				PDQLVNRSIQQGFCFNILCVGETGIGKSTLIDT
	.					LFNTNFEDYESSHFCPNVKLKAQTYELQESN
i		l	1	i i		VQLKLTIVNTVGFGDQINKEERQLGRSQSTEN
1		ł				PQKYRSEQHPVEPKKCTSFWKGALGKWAGIE
		Į			}	SSGQSAQQPYLPINSPPHRLADVADVHLFSSV
		1		. <b>.</b>	i	LSGAPGCYHLDVTVNEFKKQQNRDEQEGYS
						KGDQEQGSWKHGADPLRGGEM
1278	2628	A	10036	3	457	RAFDVRRKKSLRPCCPRDFHAGCLTVSGPST
		l		<b>.</b>		VMGAVGESLSVQCRYEEKYKTFNKYWCRQP
		}		i	, J	CLPIWHEMVETGGSEGVVRSDQVIITDHPGDL
		J	. ]			TFTVTLENLTADDAGKYRCGIATILQEDGLSG
		<b> </b>	l	. 1		FLPDPFFQVQVLVSSASSTENSVKTP
1279	2629	A	10039	214	435	NDSLVPMSSWRSCARAPSSESAWRRSAATRR
		1			1	SRKCLRTKRKRWSSGKGTQMQSTLSETPRRA
		ı		J		QMPCMWWYPFWG
1280	2630	A	10043	2	344	RATWHNAGKEREAVOLMAGAEKRVKASHS
	i i	ì	ļ	. ]	l	FLRGLFGGNTRIEEACEMYTRAANMFKMAK
	' I	i		ł	i	NWSAAGNAFCQAAKLHMQLQSKHDSATSFV
		- !	ł	ł	ľ	DAGNAYKKADPOGKTARHVACYLCV
1281	2631	A	10080	620	818	VIYKLDSSLFSYFIYFFIFETESHFLPLMKWTG
		- 1				PIMAHCSLKIL ASRNSADSAFLSAGDTSLSHST
1282	2632	A	10084	3	1640	SASIIIRGDKRASGEVGIAPSSRHILIGEPSAKY
				-	.510	NGTAIISLVRGPGILGEVTVFWRIFPPSVGEFA
	l l	ì	ł	j		ETCCKI TMDDEOCYMMUNATYMDDDDAG
	ļ	J	J			ETSGKLTMRDEQSAVIVVIQALNDDIPEEKSF
	l	i	1	1		YEFQLTAVSEGGVLSESSSTANITVVASDSPY GPEARSHEOL BUSEA ORDERTHESS GREGING
		ĺ	į	[	Í	GRFAFSHEQLRVSEAQRVNITIRSSGDFGHVR
}		l	- 1	1		LWYKTMSGTAEAGLDFVPAAGELLFEAGEM
	}		J	J		RKSLHVEILDDDYPEGPEEFSLTTTKVELQGR
						CAUGHACKICI AHAMMAA MAAMAA AMAAA
Į.			ļ	ı		GYDFTIQENGLQIDQPPEIGNISIVRIIIMKNDN AEGIIEFDPKYTAFEVEEDVGLIMIPVVRLHGT

SEQ ID NO: of nucl- cotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first amino acid residue of peptide	to last amino acid residue of peptide sequence	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
<del> </del>				sequence		nucleotide insertion YGYVTADFISQSSSASPGGVDYILHGSTVTFQ
						HGQNLSFINISIIDDNESEFEEPIEILLTGATGG AVLGRHLVSRIIIAKSDSPFGVIRFLNQSKISIA NPNSTMILSLVLERTGGLLGEIQVNWETVGPN SQEALLPQNRDIADPVSGLFYFGEGEGGVRTII LTTYPHEEIEVEETFIIKLHLVKGEAKLDSRAK
						DVTLTIQEFGDPNGVVQFAPETLSKKTYSEPL ALEGPLLITFFVRRVKGTFGEIM
1283	2633	A	10088	316	516	MGSKTLPAPVPIHPSLQLTNYSFLQAVNGLPT VPSDHLPNLYGFSALHAVHLHQWTLGYPAM
1284	2634	Α	10091	2	569	HLXRS FVSPSRAMASALIYVSKFKSFVILVVTPLLLLP
•		8		0.0		LVILMPAKFVRCAYVIILMAIYWCTEVIPLAV TSLMPVLLFPLFQILDSRQVCVQYMKDTNML
						FLGGLIVAVAVERWNLHKRIALRTILLWVGA KPARLMLGFMGVTALLSMWISNTATTAMMV
				; 		PIVEAILQQMEATSAATEAGLELVDKGKAKE LP
1285	2635	A	10092	290	728	KQSTRPDVMTLYPLHWQEEMSGESVVSSAVP
					•	AAATRTTSFKGTSPSSKYVKLNVGGALYYTT MQTLTKQDTMLKAMFSGRMEVLTDSEGWIL
						IDRCGKHFGTILNYLRDGAVPLPESRREIEELL AEAKYYLVQGLVEECQAALQV
1286	2636	Α	10100	1	574	RPRGRGAWAGPGGDYSGVRRQQRRRTRISGS
!						QRGSDAAGTMGCCTGRCSLICLCALQLVSAL ERQIFDFLGFQWAPILGNFLHIIVVILGLFGTIQ
						YRPRYIMVYTVWTALWVTWNVFIICFYLEVG
						GLSKDTDLMTFNISVHRSWWREHGPGCVRR VLPPSAHGMMDDYTYVSVTGCIVDFQYLEVI
					×	HSA
1287	2637	A	10103	252	376	RSRMGDKPIWEQIGSSFIQHYYQLFDNDRTQL GAIYVSFQL
1288	2638	A	10107	1	478	MEEEDESRGKTEESGEDRGDGPPDRDPTLSPS
	- 1					AFILRAIQQAVGSSLQGDLPNDKDGSRCHGL RWRRCRSPRSEPRSQESGGTDTATVLDMATD
			1		0.0	SFLAGLVSVLDPPDTWVPSRLDLRPGESEDM
	Ì					LELVAEVRIGDRDPIPLPVPSLLPRLRAWRTG KT
1289	2639	Α	10113	237	438	LLSRMPSTNRAGSLKDPELAELFFKEDPEKLFT DLREIGHGSFGAAYFARDVRTNEVVAIKKMS YSG
1290	2640	A	10114	367	856	RGAKAKSAVLPPGPPCSSILILSPPAPLTPRSPG
						TEATRPTAMSKSLKKKSHWTSKVHESVIGRN PEGQLGFELKGGAENGQFPYLGEVKPGKVAY
	ĺ					ESGSKLVSEELLLEVNETPVAGLTIRDVLAVI
						KHCKDPLRLKCVKQGESSGLLSVLPGGGTAR
1291	2641	A	10116	128	591	GAGQ RTIRETERRSALSCSVLKSEPLPGLQPQASQQR
						RRRLPGRRQVQVQEGGGSGLRAWVLAMASV LGSGRGSGGLSSQLKCKSKRRRRRRSKRKDK
					· .	VSILSIFLAPFKHLSPGITNTEDDDTLSTSSAE
1292	2642	A	10121	1	749	VKENRNVGNLAARPPPSGDRARGGATR
		^	.0121	*	147	QRRRFRAGLWGGHGLTDGLRRNGGCGCSAR VPRVGERLRGHRCPDPLCLLLDMLFLSFHAG
	]		ľ	j	ļ	SWESWCCCCLIPADRPWDRGQHWQLEMADT
	]		ļ			RSVHETRFEAAVKVIQSLPKNGSFQPTNEMM LKFYSFYKQATEGPCKLSRPGFWDPIGRYKW
			[		į	DAWSSLGDMTKEEAMIAYVEEMKKIIETMP
						MTEKVEELLRVIGPFYEIVEDKKSGRSSDITSD

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence		09/496	correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	İ		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1		1	amino acid residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan,
	ŀ	ļ	ļ	peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						LGNVLTSTPNAKTVNGKAESSDSGAESEEEE
1293	2643	A	10124	2	989	AC PLMSLVRVVEFVAASSAQKTPSRLENYYMVC
		''	1012	~	707	KADEKFNQLVHFLRNHKQEKHLVFFRYSSGL
		}	l			CGRGIRDSARMCSTCACVEYYGKALEVLVK
		J	ļ			GVKIMCIHGKMKYKRNKIFMEFRKLQSGILV
1			ĺ			CTDVMARGIDIPEVNWVLQYDPPSNASAFVH RCGRTARIGHGGSALVFLLPMEESYINFLAIN
-		ĺ	]	ļ		QKCPLQEMKPQRNTADLLPKLKSMALADRA
1		[				VFEKGMKAFVSYVQAYAKHECNLIFRLKDL
						DFASLARGFALLRMPKMPELRGKQFPDFVPV DVNIDTIPFKDKIREKQRQKLLEQQRREKTEN
						EGRRKFIKNKAWSKOKAKKK
1294	2644	Α	10129	91	1042	VTMYKDCIESTGDYFLLCDAEGPWGIILESLA
1			ŀ			ILGIVVTILLLLAFLFLMRKIQDCSQWNVLPTQ
						LLFLLSVLGLFGLAFAFIIELNQQTAPVRYFLF GVLFALCFSCLLAHASNLVKLVRGCVSFSWT
]						TILCIAIGCSLLQIIATEYVTLIMTRGMMFVN
		•				MTPCQLNVDFVVLLVYVLFLMALTFFVSKAT
						FCGPCENWKQHGRLIFTTVLFSIIIWVVWISML
						LRGNPQFQRQPQWDDPVVCIALVTNAWVFL LLYIVPELCILYRSCRQECPLQGNACPVTAYQ
			•			HSFQVENQELSRDKWKVLLNSDFLSHSGA
1295	2645	A	10133	376	518	RPRVVTHNSQWCFLPQDHPGWLPGQSGAPG
1296	2646	A	10135	3	551	GRGAPRQEGPGSSWRQV EWSLDPFMGIMSGQVGDLSPSQEKSLAOFRE
	20.0	••	10133	, I	331	NIQDVLSALPNPDDYFLLRWLQARSFDLQKS
1						EDMLRKHMEFRKQQDLANILAWQPPEVVRL
						YNANGICGHDGEGSPVWYHIVGSQDPKGLLL
		- 1				SASKQELLRDSFRSCELLLRECELQSQKLGKR VEKIIAIFGLEGLGLRDLWKPGIELLQE
1297	2647	A	10138	48	407	MVSSCCGSVCSDQGCGQDLCQETCCRPSCCE
1		i	Ī			TTCCRTTCCRPSCCVSSCCRPQCCOSVCCOPT
		- 1		j	-	CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRPSCCETTCCHPXCC
1298	2648	A	10156	94	453	GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG
	Í	ſ		ſ		QEMYLRFDQTTRRSPYRMSRILARHQLVTKI
		l				QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP
1299	2649	A	10161	1	393	INIVAVKNDHDFLEKDLGEPLCRRLNT PRFSELVDGRGRVSARFGGSPSKAATVRSOPT
				·	333	ASAQLENMEBAPKRVSLALOLPEHGSKDIGN
]			1			VPGNCSENPCQNGGTCVPGADAHSCDCGPGF
1	[	- 1	[		1	KGRRCELACIKVSRPCTRLFSETKAFPVWEGG
1300	2650	A	10162	98	391	VCHHV AKIASLERIMPANYTCTRPDGDNTDFRYFIYA
	l			i l		VTYTGILGPGLIGNILALWVFYGYMKETKRA
	1	ļ				VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF
1301	2651	$\overline{\mathbf{A}}$	10165	1	7545	PCIDVCITEOTCI CENI OENCEVI A MEGILICIE
		<b>1</b>	10103	•		PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG
	-	1	ļ			EKHVPGVGSARHSPQASAGGSPWQRGKAQT
J	j	]	ŀ			RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR
						LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP
					.	PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP
	ł	- 1	1		• 1	RTLSVEEPGVECNQLCLYADVTDPVLCLGOK
	j	J	j			DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP
		ļ	1			ARRLSESLHVVDENKNESKIEREHKRRTSTPV
						IMEGVQEETDTRDVKRQVERSEICTEEPQKQ

SEQ ID	SEQ ID	Met	Toro	D. 3:4.3	Day 35 A A	
NO: of	NO: of	hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide		in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ì	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	-000		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		·		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		Ì		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ĺ	(	1	peptide	Soduction	-possible nucleotide deletion.
		l		sequence		nucleotide insertion
		<del></del>	<b></b>	sequence		KSTLKNEKHLKKDDSETPHLKSLLKKEVKSS
						KEKPEREKTPSEDKLSVKHKYKGDCMHKTG
		ļ	] .			DETELHSSEKGLKVEENIQKQSQQTKLSSDDK
		ŀ				TERKSKHRNERKLSVLGKDGKPVSEYIIKTDE
			ł			NVRKENNKKERRLSAEKTKAEHKSRRSSDSK
						IQKDSLGSKQHGITLQRRSESYSEDKCDMDST
			l 1			NMDSNLKPEEVVHKEKRRTKSLLEEKLVLKS
						KSKTQGKQVKVVETELQEGATKQATTPKPD
			1			KEKNTEENDSEKQRKSKVEDKPFEETGVEPV
			l			LETASSSAHSTQKDSSHRAKLPLAKEKYKSD
						KDSTSTRLERKLSDGHKSRSLKHSSKDIKKKD
		:				ENKSDDKDGKEVDSSHEKARGNSSLMEKKL
						SRRLCENRRGSLSQEMAKGEEKLAANTLSTP
	'		,	·		SGSSLQRPKKSGDMTLIPEQEPMEIDSEPGVE
						NVFEVSKTQDNRNNNSHQDIDSENMKQKTS
				ſ	1	ATVQKDELRTCTADSKATAPAYKPGRGTGV
						NSNSEKHADHRSTLTKKMHIQSAVSKMNPGE
					(	KEPIHRGTTEVNIDSETVHRMLLSAPSENDRV
1			. 1			QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS
			1		ł	LSSVTVVPLRESYDPDVIPLFDKRTVLEGSTA
	ı			i		STSPADHSALPNQSLTVRESEVLKTSDSKEGG
1						EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ
			l l			GKVIMPLGSKLTGVIVENENTIKEGGLVDMA
1			١ ، )	- 1		KKENDLNAEPNLKQTIKATVENGKKDGIAVD
			ļ		į	HVVGLNTEKYAETVKLKHKRSPGKVKDISID
				i		VERRNENSEVDTSAGSGSAPSVLHQRNGQTE
. ]	j			J	j	DVATGPRRAEKTSVATSTEGKDKDVTLSPVK
•					· l	AGPATITSSETRQSEVALPCTSIEADEGLIIGT
	i					HSRNNPLHVGAEASECTVFAAAEEGGAVVTE
	. [		ĺ	1	ľ	GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE
. [	1	- 1		Į.		KCDGSLSRDSEIVEGTTTFISEVESDGAVTSAG
i	ľ		- 1	ľ	1	TEIRAGSISSEEVDGSQGNMMRMGPKKETEG
					1	TVTCTGAEGRSDNFVICSVTGAGPREERMVT
- 1	1	- 1		l	1	GAGVVLGDNDAPPGTSASQEGDGSVNDGTE
- 1	1	ı			1	GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS
ļ		j	J	.	j	SESEENGESAMDSTVAKEGTNVPLVAAGPCD
i	l					DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH
,		ŀ		.	1	ASTCTGLGEESEGVLICESAEGDSOIGTVVEH
ſ	Į.	ſ	1	l	1	VEAEAGAAIMNANENNVDSMSGTEKGSKDT
1	i	1	1	•	ł	DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE
- 1	1	l	i	ľ		<b>GPMTSAASDQSDSQLEKVEDTTISTGLVGGS</b>
1	1	- 1	ļ	,	Ì	YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE
	. [	İ	ŀ		,	NEECDGLMATTASGDITNQNSLAGGKNQGK
ľ	ľ	ľ	1	1	1	VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE
	ŀ	ł	ļ	- 1	j	ENMEGTRYTTEEFEAPMPSAVSGDDSQLTAS
- 1	Į	- 1	1	1	}	RSEEKDECAMISTSIGEEFELPISSATTIKCAES
I	I	ļ	İ	1	1	LQPVAAAVEERATGPVLISTADFEGPMPSAPP
. 1	ļ	]	j	]	J	EAESPLASTSKEEKDECALISTSIAEECEASVS
j			ļ			GVVVESENERAGTVMEEKDGSGIISTSSVEDC
ļ	ŀ	ļ	- 1	1		EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS
ĺ		ĺ	- 1	- 1	ľ	TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA
ļ	1			1		ALISTSTAECMPISASIDRHEENQLTADNPEGN
Ì	ł		ł	- 1	İ	GDLSATEVSKHKVPMPSLIAENNCRCPGPVR
,		1		l	- 1	GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH
- 1	]	1	}	ł	1	PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL
	1	ł	- 1		1	HLINAEEKNVLLNSLQKEDKSPETGTAGGSST
- 1	ł		ļ	ı		ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK
			1		i i	DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA
1	1	1	, i	1		EHSFLPAEQQGSEDNLKTSTTKCTTGQESKIAP

NO. of   No. of   N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SOS   Sequence	NO: of	NO: of	hod	ID NO:	beginning		
Sequence			1		L.		
1914   ng to first amino acid residue of peptide peptide peptide of peptide							
mino scid residue of peptide sequence   peptide sequence   peptide sequence   peptide sequence   peptide sequence   peptide sequence   peptide sequence   pertise			1				
Persidue of peptide sequence		l	1	1			
	]	j	J	]	residue of		
SITIMEPPATYSVALLAPECCEULTEKNDYSGE   WITOASAEKTGIDINSTKSFPEEGDIMTYYS   SERNYCDIGNESPI, NVI. GOLLIK ANILKMEA, A					peptide		
WITDQASARKTODINTRISPREGRIDMYTVS   SERANCOIGNRESPLANLOGICK, KANLKMPA   YVPSEIEKNOEILAPPESLCOGKPSGIAELQR   PLAVINESLAVENSGERSTNAKGEISS   GRKDNAGALISGISVEADPKEVEEERIIMPKR   KRKQHYLSSEDPDDPDVDLSGESSYNKGEISS   GRKDNAGALISGISVEADPKEVEEERIIMPKR   KRKQHYLSSEDPDDPDVDLSGESTAQRC   PETEPHATKEENSGILEELPKTSSEINSTTSRS   MEEKDEYSSETTCEKPEONDDDTIKSGE   PETEPHATKEENSGILEELPKTSSEINSTTSRS   MEEKDEYSSETTCEKPEONDDDTIKSGE   PETEPHATKEENSGILEELPKTSSEINSTTSRS   MEEKDEYSSESTTCEKPEONDDTIKSGE   PETEPHATKEENSGILEELPKTSSEINSTTSRS   MEEKDEYSSESTTCEKPEONDDTIKSGE   PETEPHATKEENSGILEELPKTSSEINSTTSRS   MEEKDEYSSESTTCEKPEONDDTIKSGE   PETEPHATKEENSGILEELPKTSSEINSTTSRS   MEEKDEYSSESTTCEKPEONDDTIKSGE   PETEPHATKEENSGILEELPKTSSEINSTTSRS   MEEKDEYSSESTTCEKPEONDTIKSGE   PETEPHATKEENSGILEELPKTSSEINSTTSRS   MEEKDEYSSESTTALEEPTHLPLIADILEELPKTSSEINSTTSRS   MEEKDEYSSESTTALEEPTHLPLIADILEELPKTSSEINSTTSRS   MEEKDEYSSESTTALEEPTHLPLIADILEELPKTSSEINSTTSRS   MEEKDEYSSESTTALEEPTHLPLIADILEELPKTSSEINSTTSRS   MEEKDEYSSESTTALEEPTHLGESTSCHAFTS   MEEKDEYSSESTTALEEPTHLGETSTALEEPTHLGE   MEEKDEYSSESTTAL		<u></u>			sequence		
							SHTMIPPATYSVALLAPKCEQDLTIKNDYSGK
1302   2652   A   10167   321   842   EFSLIPPILAPPENLOGKENSIALELER   FULVINESLAVENSEGENIAMERISSYNKGEISS   GRKDNABAINGHSVEADPRK VEEEREIMPRIC   RKRQFYL SSEDEPDDNIPPULOSBETAQROC   PETEPHATICENSRÜLELI PKTSSEINSTISR   MEEKBYSSSEITIGEK PEONDOTHIKSQB   EFSLIPPILREPSRÜLELI PKTSSEINSTISR   MEEKBYSSSEITIGEK PEONDOTHIKSQB   FVFYFELS VIEPPKEL AKVETMEBQ VILLERGE   NSITV AGRGISTSKELSPEPPL PLILALILLE   DOFOCHIP FIGL. VALEV MEEBQ VILLERGE   NSITV AGRGISTSKELSPEPPL PLILALILLE   DOFOCHIP FIGL. VALEV MEEBQ VILLERGE   NSITV AGRGISTSKELSPEPPL PLILALILLE   DOFOCHIP FIGL. VALEV MEEBVETENDE   RKNSQQVFKLIKKK   NSITV AGRGISTSKELSPEPPL PLILALILLE   DOFOCHIP FIGL. VALEV MEEBVETENDE   RKNSQQVFKLIKKK   NSITV AGRGISTSKELSPEPPL PLILALILLE   DOFOCHIP FIGL. VALEV MEEBVETENDE   RKNSQQVFKLIKKK   NSITV AGRGIST VALEV MEEBVETENDE   RKNSQQVFKLIKKK   NSITV AGRGIST VALEV MEEBVETENDE   RKNSQQVFKLIKKK   ALMGI QAFREAA ADAQVALTLERG WEEBVETENDE   RESCLIHLIT QGGRGGICAPPLSGALQEPILGK   ALMGI QAFREAA AVGOTLE GGGRANGA FOR FIGL. VALEV MEEBVETENDE   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPLSGALQEPILA   RESCLIHLIT QGGRGGICAPPL QGGRGGICAPPL   RESCLIHLIT QGGRGGICAPPL   RESCLIHLIT QGGRGGICAPPL   RESCLIHLIT QGGRGGICAPPL   RESCLIHLIT QGGRGGICAPPL   RESCLIHLIT QGGRGGICAPPL   RESCLIH QGGRGGICAPPL   RESCLIHLIT QGGRGGICAPPL   RESCLIHLIT QGGRGGICAPPL   RESCLIHLIT QGGRGGICAPPL   RESCLIHLIT QGGRGGICAPPL   RESCLIHLIT QGGRGGICAPPL   RESCLIHLIT QGGRGGICAPPL   RESCLIH QGGRGGICAPPL   RESCLIH QGGRGGICAPPL   RESCLIH QGGRGGICAPPL   RESCLIH QGGRGGICAPPL   RESCLIH QGGRGGICAPPL   RESCLIH QGGRGGICAPPL   RESCLI							
1302   2652   A   10167   321   842   EPSI-PPILROPERVEERERI-MPER	[	Ĭ		1	ĺ		
GRRDMARAISCHSVEADPKEVEEERIMPRR   KRRQHYLSEEDPDDPWDV DIETAQRQC   PETEPHATKEENSRDLEPLYSTSTINSTISRY   MEEKDEYSSSTTIGKPERONDDTIKSQE   PETEPHATKEENSRDLEPLYSTSTINSTISRY   MEEKDEYSSSTTIGKPERONDDTIKSQE   PETEPHATKEENSRDLEPLYSTSTINSTISRY   MEEKDEYSSSTTIGKPERONDDTIKSQE   PETEPHATKEENSRDLEPLYSTSTINSTISRY   MEEKDEYSSSTTIGKPERONDDTIKSQE   PETEPHATKEENSRDLEPLYLLADLED   DETEPHYRICLAVEPKEHWTEGONFSPFP   LAND   LA			1		1		
1302   2652   A   10167   321   842   EFSLEPPLATKENSNLEELPFENSSNISTISST	1	[	1		1		
PETEPHATKEENSRILEELEKTSSETISTISRY   MEEKDEYSSETIGEKPEONDOTIKSQE   MEKEDEYSSETIGEKPEONDOTIKSQE   PESLEPELEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKEKTSSETIGEKPEONDOTIKSQE   PESLEPELEKEKTSCHEREKTSCHEREKTSCHEREKTSGE   PESLEPELEKEKTSCHEREKTSCHEREKTSGE   PESLEPELEKEKTSCHEREKTSCHEREKTSGE   PESLEPELEKEKTSCHEREKTSCHE			l	l			
1302   2652   A   10167   321   842   SESTPTEREPREPREPREPREPRE AGPEPH FVFYFTS.YVHPPKEL AGPEPH FVFYFTS.YVHPPKEL AKYEVMEBOVILTEG NSTVAGRGTSVRCLSPSPREPREPREPREPREPREPREPREPREPREPREPREPREP	i		1	ł	ł		
1302   2652   A   10167   321   842     EPSLFPILPSPARPPRPPAFFSPELAGPEPH   FVFYFILSYVPHEQUYLTEKG   NSTVAGRGISVRCLSFSRP.PPLPILALLE DGGGEHPPYHCHLAGVYREHWTPEGNPSPFP   EARETKCYVRSSVGCVEPLITQAGVTENLDR   KNSQVYFILLKKK   NSTVAGRGISVRCLSFSRP.PPLPILALLE DGGGEHPPYHCHLAGVYREHWTPEGNPSPFP   EARETKCYVRSSVGCVEPLITQAGVTENLDR   KNSQVYFILLKKK   SMSQVYFILLKKK   SMSQVYFILLKKK   SMSQVYFILLKKK   SMSQVYFILLKKK   SMSQVYFILLKKK   SMSQVYFILLKKK   SMSQVYFILLKKK   SMSQVYFILLKKK   SMSQVYFILLKKK   SMSQVYFILLKKK   LSQEDTEVKCENVTVUDKARDLLDSVTRKGA   RACEICITY   LSQEDTEVKCENVTVUDKARDLLDSVTRKGA   RACEICITY   LSQEDTEVKCENVTVUDKARDLLDSVTRKGA   RACEICITY   LSQEDTEVKCENVTVUDKARDLLDSVTRKGA   RACEICITY   LSQEDTEVKCENVTVUDKARDLLDSVTRKGA   RACEICITY   LSQEDTEVKCENVTVUDKARDLLDSVTRKGA   RACEICITY   LSQEDTEVKCENVTVUDKARDLLDSVTRKGA   RACEICITY   LSQEDTEVKCENVTVUDKARDLLDSVTRKGA   RACEICITY   LSQEDTEVKCENVTVUDKARDLLSVTRKGA   RACEICITY   LSQEDTEVKCENVTVUDKARDLLSVTRKGA   RACEICITY   LSQEDTEVKCENVTVUTKARDLLSVTRKGA   RACEICITY   LSQEDTEVKCENVTVUTKARDLAGVER   LSQEDTAMALAGVALTUVAMYPPYPPHR   DSSVTLSFGESTYNSKSVERKSCENKENKUKQL   SNSLQRNMILFILAFILFCGLLFYINLADHWKG   RNTCT   LSGEDTAMALAGTILAFILFCGLLFYINLADHWKG   RNTCT   LSGEDTSTAVKLGCSFSKENGOVER   LSGEDTSTAVKLGCSFSKENGOVER   LSGEDTSTAVKLGCSFSKENGOVER   LSGEDTSTAVKLGCSFSKENGOVER   LSGEDTSTAVKLGCSFSKENGOPE   LSGEDTSTAVKLGCSFSKENGOPE   LSGEDTSTAVKLGCSFSKENGOPE   LSGEDTSTAVKLGCSFSKENGOPE   LSGEDTSTAVKLGCSFSKENGOPE   LSGEDTSTAVKLGCSFSKENGOPE   LSGEDTSTAVKLGCSFSKENGOPE   LSGEDTSTAVKLGCSFSKENGOPE   LSGEDTSTAVKLGCSFSKENGOPE   LSGEDTSTAVKLGCSFSKENGOPE   LSGEDTSTAVKLGCSFSKENGOPE   LSGERTFORFOLOR   LSGERT							
1303	1302	2652	Α	10167	321	842	
NSTVAGRGTSVRCLSPSPRPLPPLIJADLLE	ĺ		1.	·			FVFYFFLSYVHPPKELAKYEYMEEOVILTEKG
BARTKCTVRSSVGCVEPLTTQAEVTENLDR   KNSQQVFALLKKK   KNSQQVFALLKKK   LSQEDTEIVKCENTVUDKARDLDSVTRKGA   A 10184   970   1524   LCTLSFGISGTAGSCTTEGGLDELLETNV   LSQEDTEIVKCENTVUDKARDLDSVTRKGA   A 10184   970   1524   LCTLSFGISGTAGSCTTEFGTELGTSFAQNG   ACEICITY				}			NSTVAGRGTSVRCLSPSPRPLPPLLPLLADLLE
1303   2653   A   10171   206   429   NMILLKERILINSIGEGTINGILDELLETNV   LSQEDTEIVKCENVTVIDKARDILDSVIRKGA   RACEICITYI   LSQEDTEIVKCENVTVIDKARDILDSVIRKGA   RACEICITYI   LSQEDTEIVKCENVTVIDKARDILDSVIRKGA   RACEICITYI   LCTLSFGISGTAGSCLTTEPGTELGTSFAQNGF   YHEAVVLPTQALKINPQDHRILGORSSCHER   LGQPAWALADAQVALTLRPGWPRGIFRIGK   ALMGLQRFREAAAVPQETILRGGSOPDAAREL   RSCLLHLTIQGQRGGICAPPLSPGALQPLPHA   ELAPSGLFSLRCPRSTALRSPGLSPLH   BLAPSGLFSLRCPRSTALRSPGLSPLH   DFISVTLSFGES YDNSKSWRRSGWRKWKQL   SRIVENWHILFILAFLLFGGLLFVINLADHWKG   IRNTCT   SPESTISLEGPLSKWTNVMKGWQYRWFVLDY   NAGILSTYTSKDKMMRGSRRGCVRLRGAVI   GIDDEDDSTFITTVDQKTFHFQARDADEREK   WHALEETILRHTIQLQVRVFTWFPDSSLVGA   FFFWLVSGFFFK   GETTER   SPPQNYTISSD   SPPQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPQNYTISSD   SPPQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPQNYTISSD   SPPQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQTEVQLS   SPPCQNYTISSD   SPPCQNYTISSD   SPPLSFPLDISQLQPPLPQVVIKTQT   SPPLSFPLDISQLQPPLPQVVIKTQT   SPPLSFPLDISQLQPPLPQVVICTQ   SPPLSFPLDISQLQPPLPQVVICTQ   SPP	1		ì	i :			DGFGEHPFYHCLVAEVPKEHWTPEGNPSPFP
1303   2653   A   10171   206   429   NMILLKERRILINSIGEGTINGILDELLETINV   LSQEDTEIVKCENVTVIDKARDILDSVIRKGA   RACEICITYI   1304   2654   A   10184   970   1524   LCTLSPGISGTAGSCTTEPGTELGTSFAQNGF   YHEAVVLTTQALKUNPQDHRLFGNRSSCHER   LGQPAWALADAQVALTLRPGWPRGLFRLGK   ALMGLQRPREAAAVFQETLRGGSQPDAAREL   RSCLLHLTLQGQRGGCAPPLSPGALQPLPHA   ELAPSGLPSLRCPRSTALRSPGLSPLH   ELAPSGLPSLRCPRSTALRSPGLSPLH   ELAPSGLPSLRCPRSTALRSPGLSPLH   DFLSVTLSFGESYDNSKSWRRSCWRKWKQL   SRLQRNMILFLLAFLLPCGLLFYINLADHWKG   RNTCT   IPGSTSLEGPLSWRTNVMKGWQYRWFVLDY   NAGILSYYTSKDKMRGSRRGCVRLRGAVI   GIDDEDDSTFITTVDQXTPHFQARDADERIEK   WHALESTLRHTLQLQVRVFTWFPDSSLVGA   FFFWLVSGFFFK   WHALESTLRHTLQLVRVFTWFPDSSLVGA   FFFWLVSGFFFK   WHALESTLRHTLQLVRVFTWFPDSSLVGA   FFFWLVSGFFFK   SAFPEREKVQANTVEGSPSSGRAPAATMNRVGGSPS   AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP   SAFPEREKVQANTVEGSFSGRAPAATMNRVGGSPS   AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP   SAFPEREKVQANTVEGSFSRGPLDSTRIND   PLPYRRIYAIMAAARPOSYLIRM   RAFAFFNSSWLPFHERLQVQNOGECPWQVSIQM   SRKHLCOGSILHWWVLTAAHCFRRTLLDM   AV   AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK   HKKQQSAEIQKKRIRRAFKFQRATGASLADI   MAK   LPGADVGGGHE SLRLFHLLTSAAWVPDESQ   VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ   LDFLPVKCDACQDGCCPWQVSIQM   SRKHLCOGSILHWWVLTAAHCFRRTLLDM   AV   AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK   HKKQQSAEIQKKRIRRAFKFQRATIGASLADI   MAK   LPGADVGGGHE SLRLFHLLTSAAWVPDESQ   VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ   LDFLPVKCDACQDGCCPWQVSIQM   SRCHLCOGSILHWWVLTAAHCFRRTLLDM   AV   LPGADVGGGHE SLRLFHLLTSAAWVPDESQ   VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ   LDFLPVKCDACQDGCCPWQVSIQM   DRDCDSSPGKKKKERTCRSEKTCKQ   LDFLPVKCDACQDGCCPWQVSIQM   DRDCDSSPGKKKKERTCRSEKTCKQ   LDFLPVKCDACQDGCCPWQVSIQM   DRDCDSSPGKKKKERTCRSEKTCKQ   LDFLPVKCDACQDGCDFWQVSIQM   DRDCDSSPGKKKKERTCRSEKTCKQ   LDFLPVKCDACQDGCDFWQVSIQM   DRDCDSSPGKKKKERTCRSEKTCKQ   LDFLPVKCDACQDGCDFWQVSIQM   DRDCDSSPGKKKKERTCRSEKTCKQ   LDFLPVKCDACQDGCDFWQACQDGCPWQSIQM   DRDCDSSPGKKKERTCRSEKTCKQ   LDFLPVKCDACQDGCDFWQACCCDGCPWQSIQM   DRDCDSSPGKKKERTCRSEKTCKQ   LDFLPVKCDACQDGCDFWQACCCCDGCPWQACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC							EARETKCYVRSSVGCVEPLTTQAEVTENLDR
1304   2654   A   10184   970   1524   LCTLSFGISGTAGSCLTTEPGTELGTSFAQNGF	1202	2652		10171	206		
RACEICITY    1304   2654   A   10184   970   1524   LCTLSPGISGTAGSCLTTEPGTELGTSPAQNGF	1303	2033	A	10171	206	429	
1304   2654   A   10184   970   1524	1		ł	<b>!</b>			
Second   S	1304	2654	A	10184	970	1524	
1305	1.50.		••	10104	7,0	1,52.4	
ALMGLQRRREAAAVPQETLRGGSQPDAAREL   RSCLLHLTLQQRGGICAPPLSPGALQPLPHA   RSCLLHLTLQQRGGICAPPLSPGALQPLPHA   RLAPSGLFSIRCPRSTALRSPGLSPLLH						,	
RSCLLHIT.QGQRGGICAPPL.SPGALQPL.PHA							
1305   2655   A   10194   2   394				}			
Spelttpvggapwavattvvmypppppppppppppppppppppppppppppppppp	L						
DFISVILSFGESYDNSKSWRRSCWRKWKQL   SRLQRNMILFLLAFLLFCGLLFYINLADHWKG   RNTCT   RNTCT	1305	2655	A	10194	2	394	TDLLGRRFRVDGAAMAACEGRRSGALGSSQ
SRLQRNMILFLLAFLLFCGLLFYINLADHWKG   RNTCT							
IRNTCT							DEIONING ELLAELLECCLIEVINI ADDITIO
1306		i					
NAGLLSYYTSKDKMMRGSRRGCVRLRGAVI GIDDEDDSTFITIVDQKTFHFQARDADEREK WHALESTIRHITLQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK     1307   2657	1306	2656	Α	10195	1	410	
GIDDEDDSTFTITVDQKTFHFQARDADEREK   WIHALEETILRHITLQLQVRVFTWFPDSSLVGA   FFFWLVSGFFFK	1				_		NAGLLSYYTSKDKMMRGSRRGCVRLRGAVI
1307   2657   A   10205   85   308   QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS SPDQQNYTKSR     1308   2658   A   10214   2   453   ECGGIRQPGPGPPPALASAPAATMNRVGGSPS AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTYDERTITIVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM     1309   2659   A   10233   45   421   RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRITPGSERSRWGPPLESTKENVQCGH RPAFFNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWVLTAAHCFRRTLLDM AV     1310   2660   A   10241   243   442   AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRATIGASLADI MAK     1311   2661   A   10261   751   176   LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNPCIQHRHPLDHSCRHGSRPTI KAG	1 ' 1						GIDDEDDSTFITTVDQKTFHFQARDADEREK
1307   2657   A   10205   85   308   QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS SPDQQNYTKSR     1308   2658   A   10214   2   453   ECGGIRQPGPGPPPALASAPAATMNRVGGSPS AAANYLLCTNCRKVLRKDKRRVSQPLTRGP SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM RALESRIFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTILDM AV AV AV AV AV AV AV AV AV AV AV AV AV	[ ]				·		WIHALEETILRHTLQLQVRVFTWFPDSSLVGA
DSVPLISPLDISQLOPPLPDQVVIKTQTEYQLS   SPDQQNYTKSR	1207	0655					
SPDQQNYTKSR	1307	2657	A	10205	85	308	QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM
1308 2658 A 10214 2 453 ECGGRQPGPGPPPALASAPAATMNRVGGSPS AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK  1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK  1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNPCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	] ]				j		
AAANYLLCTNCRKVLRKDKRIRVSOPLTRGP SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM  1309 2659 A 10233 45 421 RGWPEQQSTGPRDVARQPRCQKEEGRRLRP RALESRIFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV  1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK  1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHILLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEMIL QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1308	2658		10214	<del></del>	453	
SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM  1309 2659 A 10233 45 421 RGWPEQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCPRRTLLDM AV  1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK  1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	]	ا قادمت	^	10214	-	כנד	AAANVI I CTNCPVVI DVDVDDDVCODI TD CD
DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM  1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRIFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCPRRTLLDM AV  1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK  1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHILLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKHTTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	[ [	1		[	1		SAFTPEKEVVOANTVDERTNET VEEVSTSODT
PLPYRHYIAIMAAARHQCSYLINM  1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV  1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK  1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTT KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	j l						
1310 2660 A 10241 243 442 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCPRRTLLDM AV  1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK  1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTT KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	L						
RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV  1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK  1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTT KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1309	2659	Α	10233	45	421	RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP
RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV  1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK  1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VILNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	] ]		j				RALESRTFQGSERSRWGPPLESTKENVQCGH
1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK  1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHILLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP		.	٠	]			RPAFPNSSWLPFHERLQVQNGECPWQVSIQM
1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK  1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	.	ļ	1				
HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK  1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1210	200		10041			
1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1210	200U	A	10241	243	442	AFQLINAKCESAFLSKRNPLQRNWTVLYRRK
1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHILLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP			- 1	1	1		
VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1311	2661	<del>_</del> -	10261	751	176	
LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP			^	10201	,31	1/0	VII NGAICAI GEAL PAEEDDI CATICOERAGA
QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP		ļ	- 1	1		Į	
DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP		ŀ	ł	ŀ	- 1	ł	
QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG  1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP			·	j	j		DRDCDSHPGKKKEKIFTYRCSKEGCKKKEMI
KAG		l	ł	1	l	}	
5155DEG51 SAS1FWINKT OF ATSAERF VIEWF	L						
SMTILDKKDGEQAKALFEKVRKFRAHVEDSD	1312	2662	A	10270	3	669	
	L						SMTILDKKDGEQAKALFEKVRKFRAHVEDSD

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	L Amino gold engues on /A-Alesia - C-C-Ale
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1.00	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	, dance		914	ng to first	acid residue	Q=Ghtamine, R=Arginine, S=Serine,
, ucaice	{	1	7.14	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i	1	1	ĺ	peptide	sequence	
1	1	1		sequence		nucleotide insertion
<u> </u>	<del> </del>	<del> </del>	+	sequence	<del></del>	
ł	}	1	1	}	l	LIYKLYVVQTVIKTAKFIFILCYTANFVNAISF
	i					EHVCKPKVEHLIGYEVFECTHNMAYMLKKL
ł	l	1		Ì	1	LISYISIICVYGFICLYTLFWLFRIPLKEYSFEKV
	1	İ		ł		REESSFSDIPDVKNDFAFLLHMVDQYDQLYS
1313	2663	   A	10287	1221	266	KRFGVFLSEVSBNKLREISLNHEWTFEKL
1313	2005	] ^	10287	1221	200	GAHRVLSPAQGAQPRLRSAASVEVSMVGQR
ļ		ſ				VLLLVAFLLSGVLLSEAAKILTISTLGGSHYLL
j	}	1	]	}	ł	LDRVSQILQEHGHNVTMLHQSGKFLIPDIKEE
	1	1				EKSYQVIRWFSPEDHQKRIKKHFDSYIETALD
}	1	Į	1		ļ	GRKESEALVKLMEIFGTQCSYLLSRKDIMDSL
	ì	l l	1			KNENYDLVFVEAFDFCSFLIAEKLVKPFVAIL
	ł	[				PTTFGSLDFGLPSPLSYVPVFPSLLTDHMDFW
		1	1			GRVKNFLMFFSFSRSQWDMQSTFDNTIKEHF
1	i	1	1			PEGSRPVLSHLLLKAELWFVNSDCAFDFARPL
1314	2664	A	10000	636	1000	LPNTVYIGGLMEKPIKPVPQVSEPSAFSLGFT
1314	2004	A	10288	536	1890	NVQLAKFSSTLVFFFSCDADPSALAKYVLAL
i	1		1 :			VKKDKSEKELKALCIDQLDVFLQKETQIFVEK
	i	1	1			LFDAVNTKSYLPPPEQPSSGSLKVEFFPPQEK
1	l	ł .	1 1			DIKKEEITKEEEREKKFSRRLNHSPPQSSSRYR
	1					ENRSRDERKKDDRSRKRDYDRNPPRRDSYRD
ł	ł	ł	1			RYNRRGRSRSYSRSRSRSWSKERLRERDRD
İ		l	] ]			RSRTRSRSRTRSRERDLVKPKYDLDRTDPLEN
1 .	}	ĺ	1 1			NYTPVSSVPSISSGHYPVPTLSSTITVIAPTHHG
l	i	į	1 1	İ		NNTTESWSEFHEDQVDHNSYVRPPMPKKRC
·	1	ſ	i i			RDYDEKGFCMRGDMCPFDHGSDPVVVEDVN
!	<i>!</i>	l		Į		LPGMQPFPAQPPVVEGPPPPGLPPPPPILTPPPV
	1	ŀ	] [			NLRPPVPPPGPLPPSLPPVTGPPPPLPPLQPSG
	j	1	1 1	Į.		MDAPPNSATSSVPTVVTTGIHHQPPPAPPSLFT
		İ	1 1			ADTYDTDGYNPEAPSITNTSRPMYRHRVHPR
1212	000	_	10000	445	1001	AKLG
1315	2665	Α	10293	447	1331	SHPLLSCPEKVSAKLRAAAEAAAEERRTRGA
<b>[</b> ]	[	i	1 1	Í		GSRGICAGLRSVAPGPEPLKQEEGRREWGSSI
	l		[ [	ľ		GTPSPCGSAQAAAAAAAEEATEKIPALRPALL
[			1 1			WALLALWLCCATPAHALQCRDGYEPCVNEG
j l			]	1		MCVTYHNGTGYCKCPEGFLGEYCQHRDPCE
			[ [		ſ	KNRCQNGGTCVAQAMLGKATCRCASGFTGE
						DCQYSTSHPCFVSRPCLNGGTCHMLSRDTYE
]			1			CTCQVGFTGRNPKCPGGNLNYQFNGIIVVYS
1 1				ì		GGSVPPSGTKTSKPAEHNAMGTGSKNFASGT
						LWVMVSGATSTSTSTL
1316	2666	A	10294	118	572	SLSMESNHKSGDGLSGTQKEAALRALVQRTG
				. 1	1	YSLVQENGQRKYGGPPPGWDAAPPERGCEIFI
				ĺ		GKLPRDLFEDELIPLCEKIGKIYEMRMMMDF
				}	J	NGNNRGYAFVTFSNKVEAKNAIKQLNNYEIR
		1				NGRLLGVCASVDNCRLFVGGIPKTKK
1317	2667	Α	10301	158	1956	LLKSCGVLLSGVCIPCEGKGPTVLVIQTAVPQ
						DRPTKSSMRSAAKPWNPAIRAGGHGPDRVRP
		Į		Į.	ļ	LPAASSGMKSSKSSTSLAFESRLSRLKRASSE
		j			1	DTLNKPGSTAASGVVRLKKTATAGAISELTES
		- 1	İ	1	İ	RLRSGTGAFTTTKRTGIPAPREFSVTVSRERSV
					l	PRGPSNPRKSVSSPTSSNTPTPTKHLRTPSTKP
		i	1	- 1	í	KQENEGGEKAALESQVRELLAEAKAKDSEIN
}	<b></b>	ľ	' [	i	ĺ	RLRSELKKYKEKRTLNAEGTDALGPNVDGTS
1	•	•	ĺ	1	ſ	VSPGDTEPMIRALEEKNKNFQKELSDLEEENR
		ĺ	ļ		· . ]	VLKEKLIYLEHSPNSEGAASHTGDSSCPTSITO
İ	ì	İ		1	·j	ESSFGSPTGNQLSSDIDEYKKNIHGNALRTSG
	ŀ	l	- 1	J	}	SSSSDVTKASLSPDASDFEHITAETPSRPLSSTS
ļ	l	1	1	1		NPFKSSKCSTAGSSPNSVSELSLASLTEKIOKM
İ	ì	ł	1	1	1	EENHHSTAEELQATLQELSDQQQMVQELTAE
						, -, (440.v. Apply

SEQ ID	SEQ ID	Met	SEQ	D-3:4-3	18 0	
NO: of	NO: of	hod	ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	110th	in NO.	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutarnic Acid,
eotide	seq-	ļ	USSN	location		F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence		09/496	correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine,
uence	00.00	1	914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline,
1	Ĭ	l	717	amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine,
1	ł	ļ	l	residue of	sequence	T=Threonine, V=Valine, W=Tryptophan,
1	l	1	1	peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
		l	l	sequence		nucleotide insertion
<b></b>	<del> </del>	<del> </del>	<del> </del>	Soquelice	<del> </del>	MENT VIDENTIL PROFITOUR PROFITOUR A DOLL CORNEY
1	1	Į.		1		NEKLVDEKTILETSFHQHRERAEQLSQENEKL
		i .	1		ł .	MNLLQERVKNEEPTTQEGKIIELEQKCTGILE
i	1					QGRFEREKLLNIQQQLTCSLRKVEEENQGAL
ł	ļ	l		ſ		EMIKRLKEENEKLNEFLELERHNNNMMAKTL EECRVTLEGLKMENGSLKSHLQG
1318	2668	A	10303	333	879	
-2.5	2000	1 ^	10303	333	0/3	GECFIMAAVVQQNDLVFEFASNVMEDERQL
i.	ŀ					GDPAIFPAVIVEHVPGADILNSYAGLACVEEP
1 .	j	ļ	]	j	ļ	NDMITESSLDVAEEEIIDDDDDDDTTLTVEASCH
	l		i		ŀ	DGDETIETIEAAEALLNMDSPGPMLDEKRINN NIFSSPEDDMVVAPVTHVSVTLDGIPEVMETQ
l	j					UNDERAY DEBOY CEDE OPADARA
1319	2669	A	10322	169	654	QVQEKYADSPGASSPEQPKRKKK
		1 **	10322	107	357	MEVRMSGSVAVTRAIAVPGLLLLIIATALSL
		l			•	LIGAKSLPASVVLEAFSGTCQSADCTIVLDAR
[		Ì	1		Ì	LPRTLAGLLAGGALGLAGALMQTLTRNPLAD PGLLGVNAGASFAIVLGAALFGYSSAQEQLA
1			<b>!</b>			MARACAL WASI TVASTOSOCOCOLORUM
1			Ì			MAFAGALVASLIVAFTGSQGGGQLSPVRLTL AGVXL
1320	2670	Α -	10323	441	2	KMNQVAVVIGGGQTLGAFLCHGLAAEGYRV
1.520	20,0	**	10323	441	2	
{	:		[			AVVDIQSDKAANVAQEINAEYGESMAYGFG ADATSEQSVLALSRGVDEIFGRVDLLVYSAGI
						AKAAFISDFQLGDFDRSLQVNLVGYFLCARE
(						
1321	2671	Α	10332	1	453	FSRLMIRDGIQGRIIQINSKSDE RHRTAGPGSTISSRTDSASAPAARAMPCEYTY
-5	5071	' '	10332	•	433	AKLTSDCSRPSLQWYTRAQSKMRRPRLLLKD
						ILKCTLLVFGVRILYILKLNYTTEECDMKNMH
i i						YVDPDHVKRAQKYAQQVLQKESPPKFAKTS
						MALLFEHRYSVDLLPFVQKAPTDSEA
1322	2672	A	10333	25	423	EPSNGPVVYSALGNEDDEILLLGKDIIGTFAAS
			10000			ERKMRAHQVLTFLLLFVITSGASENASTSRGC
í I						GLDLLPQNVYLCDLDAIWGIVVEAVAGAGA
1 I		i				LITLLLMLILLGRLPFIKEKEKKSPAVLHFLFL
1 1				i		LGTLG
1323	2673	Α	10334	52	426	SSLGNEDDEILSLAKDITGMFVASHRKMRAH
						QVLTFLLLFVITSVASENASTSRGCGLDLLPQ
				ļ	ļ	YVSLCDLDAIWGIVVBAAAGAGALITLLLMLI
				į.	·	LLVRLPFFKEKEKKSPVGLHFLFLLGTLGP
1324	2674	Α	10336	i	932	ERLCFPCMQSKIYSYMSPNKCSGMRFPLQEE
	ſ	,		ľ		NSVTHHEVKCQGKPLAGIYRKREEKRNAGN
		i				AVRSAMKSEEOKIKDARKGPLVPFPNOKSEA
	ł			1	.	AEPPKTPPSSCDSTNAAIAKQALKKPIKGKQA
]				-		PRKKAQGKTQQNRKLTDFYPVRRSSRKSKAE
			]	. 1	]	LQSEERKRIDELIESGKEEGMKIDLIDGKGRG
	. 1	1	}	ļ	J	VIATKQFSRGDFVVEYHGDLIEITDAKKREAL
	ļ			ļ	1	YAQDPSTGCYMYYFQYLSKTYCVDATRETN
	j			į	ì	RLGRLINHSKCGNCQTKLHDIDGVPHLILIAS
		- 1	1	1	ľ	RDIAAGEELLYDYGDRSKASIEAHPWLKH
1325	2675	A	10338	3	870	PGSTISCSELKGTQCRATAGSRGRRPPMTCWL
	ı	I		l	ł	RGVTATFGRPAEWPGYLSHLCGRSAAMDLG
			l		]	PMRKSYRGDREAFEETHLTSLDPVKQFAAWF
	ļ		J	j	j	EEAVQCPDIGEANAMCLATCTRDGKPSARML
· .			l		İ	LLKGFGKDGFRFFTNFESRKGKELDSNPFASL
	}	ļ	[	l	ł	VFYWEPLNRQVRVEGPVKKLPEEEAECYFHS
				ĺ	1	RPKSSQIGAVVSHQSSVIPDREYLRKKNEELE
	. !	1	- 1	ĺ	ſ	QLYQDQEVPKPKSWGGYVLYPQVMEFWQG
ĺ	ľ			1	. 1	QTNRLHDRIVFRRGLPTGDSPLGPMTHRGEE
		_	j	ì	- 1	DWLYERLAP
1326	2676	A	10344	2	984	ARAAAHCGICRLVRWWRKRRSVMGIQTSPV
	1	ł		J		LLASLGVGLVILLGLAVGSYLVRRSRRPOVT
		1			. 1	LLDPNEKYLLRLLDKTTVSHNTKRFRFALPTA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, I=possible nucleotide deletion, V=possible
				sequence		nucleotide insertion  HHTLGLPVGKHIYLSTRIDGSLVIRPYTPVTSD EDQGYVDLVIKVYLKGVHPKFPEGGKMSQY LDSLKVGDVVEFRGPSGLLTYTGKGHFNIQP NKKSPPEPRVAKKLGMIAGGTGITPMLQLIRA ILKVPEDPTQCFLLFANQTEKDIILREDLEELQ ARYPNRFKLWFTLDHPPKDWAYSKGFVTAD MIREHLPAPGDDVLVLLCGPPPMVQLACHPN
1327	2677	A	10345	1	968	LDKLGYSQKMRFTY  LQSAGEGVTHVLILLESPARPVAAVTQVQRR RYHRLSDMSMLAERRKQKWAVDPQNTAW SNDDSKFGQRMLEKMGWSKGKGLGAQEQG ATDHIKVQVKNNHLGLGATINNEDNWIAHQ DDFNQLLAELNTCHQETTDSSDKKEKKSFS LEEKSKISKNRVHYMKFTKGKDLSSRSKTDL DCIFGKRQSKKTPEGDASPSTPEENETTTTSAF TIQEYFAKRMAALKNKPQVPVPGSDISETQVE RKRGKKRNKEATGKDVESYLQPKAKRHTEG KPERAEAQERVAKKKSAPAEEQLRGPCWDQ SSKASAQDAGDHVQPA
1328	2678	A	10346	173	439	GSAAMKVKIKCWNGVATWLWVANDENCGI CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF
1329	2679	A	10351	3	964	HMHCILKWLHAQQVQQHCPMCRQEWKFKE QMEPGNDTQISEFLLLGFSQEPGLQPFLFGLFL SMYLVTVLGNLLIILATISDSHLHTPMYFFLSN LSFADICVTSTTIPKMLMNIQTQNKVITYIACL MQMYFFILFAGFENFLLSVMAYDRFVAICHP LHYMVIMNPHLCGLLVLASWTMSALYSLLQI LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF LNHMVIYFTVALLGGGPLTGILYSYSKISSIH AISSAQGKYKAFSTCASHLSVVSLFYGAILGV YLSSAATRNSHSSATASVMYTVVTFMLNPFI YSLRNKDIKRALGIHLLWGTMKGQFFKKCP
1330	2680	A	10352	34	2573	IPFLKSCCCCLFDFPPPPLDQVQEEECEVERV TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRITLKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGLLKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKHEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE

SEQ ID NO: of NO: of nucleotide eotide sequence With the nucleotide uence where the nucleotide nucleotide sequence where the nucleotide uence where the nucleotide sequence where the nucleotide sequence where the nucleotide sequence where the nucleotide location corresponding to last amino acid residue of peptide sequence where the nucleotide location corresponding to last amino acid residue of peptide sequence where the nucleotide location corresponding to last amino acid residue of peptide sequence where the nucleotide deletion, the nucleotide insertion where the nucleotide deletion, the nucleotide insertion where the nucleotide deletion, the nucleotide insertion where the nucleotide insertion acid sequence (A=Alanine C=Cysteine and nucleotide in nucleotide in nucleotide, p=Aspartic Acid, E=Glutamic, Acid, E=Glutamic, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan, f=possible nucleotide deletion, the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequence where the nucleotide insertion acid residue of peptide sequ	ne,
nucleotide sequence    Deptide    Deptide sequence   Deptide s	
eotide sequence    Sequence   USSN   Ocation   Corresponding to last amino acid residue of peptide sequence   USSN   Ocation   Corresponding to last amino acid residue of peptide sequence   T=Frienyianamie, C=Grycine, H=Histodine, H=Istodeucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion    1331	
sequence    Sequence   Description	
uence  914  ng to first amino acid residue of peptide residue of peptide sequence  1331  2681  A 10353  1 2100  AVEFAEGALTMAPWPELGDAQPNPDKYI AAGQQPTAPDKSKETNKTDNTEAPVTKIP SYSTATLIDEPTEVDDPWNLPTLQDSGIKY	
amino acid residue of peptide sequence T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  1331 2681 A 10353 1 2100 AVEFAEGALTMAPWPELGDAQPNPDKYI AAGQQPTAPDKSKETNKTDNTEAPVTKIP SYSTATLIDEPTEVDDPWNLPTLQDSGIKY	
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  1331 2681 A 10353 1 2100 AVEFAEGALTMAPWPELGDAQPNPDKYI AAGQQPTAPDKSKETNKTDNTEAPVTKIP SYSTATLIDEPTEVDDPWNLPTLQDSGIKY	
peptide sequence   /=possible nucleotide deletion, \=possible nucleotide insertion   AVEFAEGALTMAPWPELGDAQPNPDKYI   AAGQQPTAPDKSKETNKTDNTEAPVTKIF   SYSTATLIDEPTEVDDPWNLPTLQDSGIK	
Sequence   nucleotide insertion   1331   2681   A   10353   1   2100   AVEFAEGALTMAPWPELGDAQPNPDKYI   AAGQQPTAPDKSKETNKTDNTEAPVTKIF   SYSTATLIDEPTEVDDPWNLPTLQDSGIKY	
1331   2681   A   10353   1   2100   AVEFAEGALTMAPWPELGDAQPNPDKYT   AAGQQPTAPDKSKETNKTDNTEAPVTKIF   SYSTATLIDEPTEVDDPWNLPTLQDSGIKY	
1331 2681 A 10353 1 2100 AVEFAEGALTMAPWPELGDAQPNPDKYT AAGQQPTAPDKSKETNKTDNTEAPVTKIF SYSTATLIDEPTEVDDPWNLPTLQDSGIKY	
AAGQQPTAPDKSKETNKTDNTEAPVTKIF SYSTATLIDEPTEVDDPWNLPTLQDSGIKY	YLEG
SYSTATLIDEPTEVDDPWNLPTLQDSGIKY	TRITE
	ZWCE.
RDTKGKILCFFQGIGRLILLLGFLYFFVCSI	Z425
SSAFQLVGGKMAGQFFSNSSIMSNPLLGL	31 744C
VLVTVLVQSSTSTSIVVSMVSSSLLTVRA	TO A TO
IIMC A NICTED THAN A CONDUCTION A	CAAIP
IIMGANIGTSITNTIVALMQVGDRSEFRRA	AMA
GATVHDFFNWLSVLVLLPVEVATHYLEIT	TITQL
IVESFHFKNGEDAPDLLKVITKPFTKLIVQ	QLDK
KVISQIAMNDEKAKNKSLVKIWCKTFTNK	NKTQ
INVTVPSTANCTSPSLCWTDGIQNWTMKN	CNVT
YKENIAKCQHIFVNFHLPDLAVGTILLILSI	SLLV
LCGCLIMIVKILGSVLKGQVATVIKKTINT	ITDFP
FPFAWLTGYLAILVGAGMTFIVQSSSVPTS	TSAL
TPLIGIGVITIERAYPLTLGSNIGTITTAILA	AAL
ASPGNALRSSLQIALCHFFFNISGILLWYPI	PIPFT
RLPIRMAKGLGNISAKYRWFAVFYLIIFFF	FLIP
LTVFGLSLAGWRVLVGVGVPVVFIIILVLC	LCLR
LLQSRCPRVLPKKLQNWNFLPLWMRSLK	KPW
DAVVSKFTGCFQMRCCCCCRVCCRACCL	TIC
GCPKCCRCSKCCEDLEEAQEGQDVPVKAI	APRT
FDNITISREAQGEVPASDSKTECTAL	ru Di
1332 2682 A 10354 30 1377 SQQGSQPHRQGPPSLLTAPHSLDLPALPPG	aasa
GSQGKLRRVLVPMSVKPSWGPGPSEGVTA	TAVE
TODI CELINIMPELLIN WHAT OF THE COMPANY OF	IAVP
TSDLGEIHNWTELLDLFNHTLSECHVELSC	SOSI
KRVVLFALYLAMFVVGLVENLLVICVNW	WRG
SGRAGLMNLYILNMAIADLGIVLSLPVWM	MLE
VILDYTWLWGSFSCRFTHYFYFVNMYSSI	SIFF
LVCLSVDRYVTLTSASPSWQRYQHRVRRA	ram
CAGIWVLSAIIPLPEVVHIQLVEGPEPMCLE	LFM
APFETYSTWALAVALSTTILGFLLPFPLETV	'VFN
VLTACRLRQPGQPKSRRHCLLLCAYVAVF	VFV
MCWLPYHVTLLLLTLHGTHISLHCHLVHL	ILLY I
FFYDVIDCFSMLHCVINPILYNFLSPHFRGR	GRLL
NAVVHYLPKDQTKAGTCASSSSCSTOHSII	smr
KGDSQPAAAAPHPEPSLSFQAHHLLPNTSP	SPISP
TQPLTPS	
1333 2683 A 10358 2 884 AAGAGADGREPASERASRAEPPAVAMGQ	OND
LMGTAEDFADQFLRVTKQYLPHVARLCLL	LIGT
FLEDGIRMWFQWSEQRDYIDTTWNCGYLI	TTA
SSFVFLNLLGQLTGCVLVLSRNFVQYACFO	ECI E
GIIALQTIAYSILWDLKFLMRNLALGGGLLI	177   177
I TARRETOR THE AUTHOR PROPERTY OF CO.	77 CC
LAESRSEGKSMFAGVPTMRESSPKQYMQL	STORY
RVLLVLMFMTLLHFDASFFSIVQNIVGTAL	TWT
LVAIGFKTKLAALTLVVWLFAINVYFNAFY	rwT
PVYKPMHDFLKYDFFQTMSVIGGLLLVVA	VAL
1334 2684 A 10367 59 1562 OAWSLOVAL SPECTRA SPENCE A AVROVAL	
1302 QAWSLQVALSFFFFASPSNSFAAAVPOLL	LFP
	ELM
ELVWGTKSSPGLSDTIFCRWTQGFVFSESE	EGS
ALEQFEGGPCAVIAPVQAFLLKKLLFSSEKS	KSS
WRDCSQEEQKELLCHTLCDILESACCDHSC	SGS
YCLVSWLRGKTTEETASISGSPAESSCQVE	EHS
SALAVEELGFERFHALIQKRSFRSLPELKDA	DAV
LDQYSMWGNKFGVLLFLYSVLLTKGIENIK	TKN
EIEDASEPLIDPVYGHGSQSLINLLLTGHAV	VON
VWDGDRECSGMKLLGIHEQAAVGFLTLME	ATC A
LRYCKVGSYLKISKIPYLDCLASETHLTVFF	
VDMAI VADE A DROW A DROWN AND	rra
KDMALVAPEAPSEQARRVFQTYDPEDNGF	IFIP
DSLLEDVMKALDLVSDPEYINLMKNKLDPI	PEG

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LGIILLGPFLQEFFPDQGSSGPESFTVYHYNGL KQSNYNEKVMYVEGTAVVMGFEDPMLQTD
1335	2685	A	10375	82	2929	DTPIKRCLQTKWPYIELLWTTDRSPSLN  TRTKRRLGREKAMASPPRGWGCGELLLPFML LGTLCEPGSGQIRYSMPEELDKGSFVGNIAKD LGLEPQELAERGVRIVSRGRTQLFALNPRSGS LVTAGRIDREELCAQSPLCVVNFNILVENKM KIYGVEVEIIDINDNFPRFRDEELKVKVNENA AAGTRLVLPFARDADVGVNSLRSYQLSSNLH FSLDVVSGTDGQKYPELVLEQPLDREKETVH DLLLTALDGGDPVLSGTTHIRVTVLDANDNA PLFIPSEYSVSVPENIPVGTRLLMLTATDPDE GINGKLTYSFRNEEEKISETFQLDSNLGEISTL QSLDYESSRFYLMEVVAQDGGALVASAKVV VTVQDVNDNAPEVILTSLTSSISEDCLPGTVIA LFSVHDGDSGENGEIACSIPRNLPFKLEKSVD NYYHLLTTRDLDREETSDYNITLTVMDHGTP PLSTESHIPLKVADVNDNPPNFPQASYSTSVT ENNPRGVSIFSVTAHDPDSGDNARVTYSLAE DTFQGAPLSSYVSINSDTGVLYALRSFDYEQL RDLQLWVTASDSGNPPLSSNVSLSLFVLDQN DNTPEILYPALPTDGSTGVELAPRSAEPGYLV TKVVAVDKDSGQNAWLSYRLLKASEPGLFA VGLHTGEVRTARALLDRDALKQSLVVAVED HGQPPLSATFTVTVAVADRIPDILADLGSIKTP IDPEDLDLTLYLVVAVAAVSCVFLAFVIVLLV LRLRRWHKSRLLQAEGSRLAGVPASHFVGV DGVRAFLQTYSHEVSLTADSRKSHLIFPQPNY ADTLLSEESCEKSEPLLMSDKVDANKEERRV
1336	2686	A .	10379	1	557	QQAPPNTDWRFSQAQRPGTSGSQNGDDTGT WPNNQFDTEMLQAMILASASEAADGSSTLGG GAGTMGLSARYGPQFTLQHVLQGELGSDYR QNVYIPGSNATLTNAAGKRDGKAPAGGNGN KKKSGKKEKK RPRRRQPSFSCRVLVLEDPPCFRFTNSMNQEK
1337	2687	A	10380	1	1263	LAKLQAQVRIGGKGTARRKKKVVHRTATAD DKKLQSSLKKLAVNNIAGIEEVNMIKDDGTVI HFNNPKVQASLSANTFAITGHAEAKPITEMLP GILSQLGADSLTSLRKLAEQFPRQVLDSKAPK PEDIDEEDDDVPDLVENFDEASKNEAN
		A			1205	IPGSTISWSPAAARGLSVCRCCRLHPASAMDL FGDLPEPERSPRPAAGKEAQKGPLLFDDLPPA SSTDSGSGGPLLFDDLPPASSGDSGSLATSISQ MVKTEGKGAKRKTSEEKNGSEELVEKKVC KASSVIFGLKGYVAERKGEREEMQDAHVILN DITEECRPPSSLITRVSYFAVFDGHGGIRASKF AAQNLHQNLIRKFPKGDVISVEKTVKRCLLD TFKHTDEEFLKQASSQKPAWKDGSTATCVLA VDNILYIANLGDSRAILCRYNEESQKHAALSL SKEHNPTQYEERMRIQKAGGNVRDGRVLGV LEVSRSIGDGQYKRCGVTSVPDIRRCQLTPND RFILLACDGLFKVFTPEEAVNFILSCLEDEKIQ TREGKSAADARYEAACNRLANKAVQRGSAD NVTVMVVRIGH
1338	2688	A	10385	3	589	GPSQSMAAGELEGGKPLSGLLNALAQDTFHG YPGITEELLRSQLYPEVPPEEFRPFLAKMRGIL KSIASADMDFNQLEAFLTAQTKKQGGITSDQ AAVISKFWKSHKTKIRESLMNQSRWNSGLRG LSWRVDGKSQSRHSAQIHTPVAIIELELGKYG QESEFLCLEFDEVKVNQILKTLSEVEESISTLIS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion OPN
1339	2689	A	10386	50	390	LGAMAKHHPDLIFCRKQAGVAIGRLCEKCDG KCVICDSYVRPCTLVRICDECNYGSYQGRCVI CGGPGVSDAYYCKECTIQEKDRDGCPKIVNL GSSKTDLFYERKKYGFKKR
1340	2690	·	10388	113	3472	SQLRKGASATHSSPSRTDCIAQMMDIYVCLK RPSWMVDNKRMRTASNFQWLLSTFILLYLM NQVNSQKKGAPHDLKCVTNNLQVWNCSWK APSGTGRGTDYEVCIENRSRSCYQLEKTSIKIP ALSHGDYEITINSLHDFGSSTSKFTLNEQNVSL IPDTPEILNLSADFSTSTLYLKWNDRGSVFPHR SNVIWEIKVLRKESMELVKLVTHNTTLNGKD TLHHWSWASDMPLECAIHFVEIRCYIDNLHFS GLEEWSDWSPVKNISWIPDSQTKVFPQDKVIL VGSDITFCCVSQEKVLSALIGHTNCPLIHLDGE NVAIKIRNISVSASSGTNVVFTTEDNIFGTVIF AGYPPDTPQQLNCEIHDLKEIICSWNPGRVTA LVGPRATSYTLVESFSGKYVRLKRAEAPTNES YQLLFQMLPNQEIYNFTLNAHNPLGRSQSTIL VNITEKVYPHTPTSFKVKDINSTAVKLSWHLP GNFAKINFLCEIEIKKSNSVQEQRNVTIKGVE NSSYLVALDKLNPYTLYTFRIRCSTETFWKW SKWSNKKQHLTTEASPSKGPDTWREWSSDG KNLIIYWKPLPINEANGKILSYNVSCSSDEETQ SLSEIPDPQHKAEIRLDKNDYIISVVAKNSVGS SPPSKIASMEIPNDDLKIEQVVGMGKGILLTW HYDPNMTCDYVIKWCNSSRSEPCLMDWRKV PSNSTETVIESDEFRPGIRYNFFLYGCRNQGY QLLRSMIGYIEELAPIVAPNFTVEDTSADSILV KWEDIPVEELRGFLRGYLFYFGKGERDTSKM RVLESGRSDIKVKNITDISQKTLRIADLQGKTS YHLVLRAYTDGGVGPEKSMYVVTKENSVGL IIAILIPVAVAVIVGVVTSILCYRKREWIKETFY PDIPNPENCKALQFQKSVCEGSSALKTLEMNP CTPNNVEVLETRSAPPKIEDTEIVSPVAERPEN RSDAKPENHVVESYCPPIEEEIPNPAADETGG TAQVIYIDVQSMYQPQAKPEEQENDPVGGA GYKPQMHLPINSTVEDIAAEEDLDKTAGYRP QANVNTWNLVSPDSPRSIDSNSEIVSFGSPCSI NSRQFLIPPKDEDSPKSNGGGWSFTNFFQNKP ND
1341	2691	A		1	5057	MLPPKHLSATKPKKSWAPNLYELDSDITKEP DVIIGEGPTDSEFFHQRFRNLIYVEFVGPRKTL IKLRNLCLDWLQPETRTKEEIIELLVLEQYLTII PEKLKPWVRAKKPENCEKLVTLLENYKEMY QPEGESLHGVLVVSAGLRCPLGLSASTLLTW SGLDNSLSWAAVGMSCVLWDIELHHDFLGV ATKSVSTHAQGDAAQGLGGTTVRMWARDSN LATGVLLDDNNSDVTSDDDMTRNRRESSPPH SVHSFSGDRDWDRRGRSRDTEPRDRWSHTR NPRSRMPPRDLSLPVVAKTSFEMDREDDRDS RAYESRSQDAESYQNVVDLAEDRKPHNTIQD NMENYRKLLSLGVQLAEDDGHSHMTQGHSS RSKRSAYPSTSRGLKTMPEAKKSTHRRGICED ESSHGVIMEKFIKDVSRSSKSGRARESSDRSQ RFPRMSDDNWKDISLNKRESVIQQRVYEGNA FRGGFRFNSTLVSRKRVLERKRRYHFDTDGK GSIHDQKGCPRKKPFECGSEMRKAMSVSSLS SLSSPSTTESQPIDFGAMPYVCDECGRSFSVIS EFVEHQIMHTRENLYEYGESFIHSVAVSEVQK

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	
eotide	seq-	1	USSN	location		F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence		09/496	1	corresponding	1=Isoleucine, K=Lysine, L=Leucine,
	ченсе	i		correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ł	ŀ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
i	1	ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	į.	]	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1		peptide		/=possible nucleotide deletion, \=possible
	ľ			sequence		nucleotide insertion
			1	<del></del>	<del>                                     </del>	SQVGGKRFECKDCGETFNKSAALAEHRKIHA
1	l	1	1		j	RGYLVECKNQECEEAFMPSPTFSELQKIYGK
1			1	ŀ		DESCRIPTION OF THE PROPERTY OF
ł	l	-	1	ł	1	DKFYECRVCKETFLHSSALIEHQKIHFGDDKD
			ĺ		1	NEREHERERERGETFRPSPALNEFQKMYG
1	ļ	l	l		]	KEKMYECKVCGETFLHSSSLKEHQKIHTRGN
ľ	ļ	1	1			PFENKGKVCEETFIPGQSLKRRQKTYNKEKLC
						DFTDGRDAFMQSSELSEHQKIHSRKNLFEGR
					1	GYEKSVIHSGPFTESQKSHTITRPLESDEDEKA
1	ł	ł	ł		l	FTISSNPYENQKIPTKENVYEAKSYERSVIHSL
1	l .		į.			ASVEAQKSHSVAGPSKPKVMAESTIQSFDAIN
1	ł		l			HQRVRAGGNTSEGREYSRSVIHSLVASKPPRS
1		[	[			HUCKEI AEGUENCEGOLAGON ALION AND AND AND AND AND AND AND AND AND AN
1	1	l	İ			HNGNELVESNEKGESSIYISDLNDKRQKIPAR
ľ			1			ENPCEGGSKNRNYEDSVIQSVFRAKPQKSVP
1 .	1	ł ·	1			GEGSGEFKKDGEFSVPSSNVREYQKARAKKK
1	1	J	<b>,</b>			YIEHRSNETSVIHSLPFGEQTFRPRGMLYECQ
1	ľ	ł i	1			ECGECFAHSSDLTEHQKIHDREKPSGSRNYE
1			•			WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK
						DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE
						NTDGEETHSEETHGQETIEDPVIQGSDMEDPQ
						KDDPDDKIYECEDCGLGFVDLTDLTDHQKVH
	[					SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ
1						LYECPKCGESFIHSSFLFEHQRIHEQDQLYSM
l i						RCCDDCEIVI I DVANDDDDD V VEDVIDVI V VOV
i l						KGCDDGFIALLPMKPRRNRAAERNPALAGSA
						IRCLLCGQGFIHSSALNEHMRLHREDDLLEQS
ł l						QMAEEAIIPGLALTEFQRSQTEERLFECAVCG
j l						ESFVNPAELADHVTVHKNEPYEYGSSYTHTS
			ĺ			FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE
i i						LHLEEEEEDEAAAAAAAAQEVEANVHVPQ
						VVLRIQGLNVEAAEPEVEAAEPEV
1	:		•			EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE
ł I			l			QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE
}				J	J	GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC
1						HECTETFTSSTAFSEHLKTHASMIIFEPANAFG
j i						ECSGYIERASTSTGGANQADEKYFKCDVCGQ
1						
1342	2692	A	10393	2	1350	LFNDHLSLARHQNTHTG
	2072	^	10373	- 1	1330	GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA
, ,		1				ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS
			•			APAVLVVAVAVVVVVVSAVAWAMANYIHV
1 1		- 1		ļ	1	PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ
		]				HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN
į l		J		J		TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL
j l				ĺ	].	QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH
		I				VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL
Į į		1			ļ	WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI
; )		ŀ	. ]	1	1	
j l			Ì		ŀ	LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII
1 1	ł	ŀ	ì			YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE
1	1			1	Ì	FAGVSDVDYSLYPDRELQSQWLRAYLEAYK
	j	j	· }	1	J	EFKGFGTEVTEKEVEILFIQVNQFALASHFFW
		l		l	ĺ	GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM
<b>├</b>						KPEVTALKVPE
1343	2693	A	10394	102	839	PEAQTSAVLAREKGHLPTMRHEAPMQMASA
1	1	. }				QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK
				j	ļ	TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN
	1	1			į	ENDINI OLIMATA COMPANIMENTA MOTO TO COMP
j l			ł	İ	•	EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI
1		1	. 1	•		LMYDITNEESFNAVQDWSTQIKTYSWDNAQ
1		Į				VILVGNKCDMEDERVISTERGQHLGEQLGFE
1	ŀ	ŀ	ł	1	· 1	FFETSAKDNINVKQTFERLVDIICDKMSESLET
	l					DPAITAAKQNIRLKETPPPPQPNCAC
1344	2694	A	10395	2	4136	DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS
	l	ŀ		J	·	LQQRTPAEMSPVLHFYVRPSGHEGAASGHTR
						TIMOUNDUINK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Day diag a series	
NO: of	NO: of	hod	ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	]	in	nucleotide	location	D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l .	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	Ì	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		Į		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
	1	ĺ		peptide		/=possible nucleotide deletion, \=possible
•	L		i	sequence		nucleotide insertion
						RKLQGKLPELQGVETELCYNVNWTAEALPS.
	1	l .	1	ì		EETKKLMWLFGCPLLLDDVARESWLLPGSN
			1		Ì	DLLLEVGPRLNFSTPTSTNIVSVCRATGLGPV
	İ			1	İ	DRVETTRRYRLSFAHPPSAEVEAIALATLHDI
		J				MTEQHFPHPIQSFSPESMPEPLNGPINILGEGR
	† ·	ļ		1		LALEKANQELGLALDSWDLDFYTKRFQELQI
	l					NPSTVEAFDLAQSNSEHSRHWFFKGQLHVDC
	l		1			QKLVHSLFESIMSTQESSNPNNVLKFCDNSSA
		[	1		ĺ	IQGKEVRFLRPEDPTRPSRFQQQQGLRHVVF
						AETHNFPTGVCPFSGATTGTGGRIRDVQCTG
		l	1	1		RGAHVVAGTAGYCFGNLHIPGYNLPWEDLS
		1		l		QYPGNFARPLEVAIEASNGASDYGNKFGEPV
		l		1		LAGFARSLGLQLPDGQRREWIKPIMFSGGIGS
		1	1	ļ		MEADHISKEAPEPGMEVVKVGGPVYRIGVGG
			J	ł		GAASSVQVQGDNTSDLDFGAVQRGDPEMEQ KMNRVIRACVEAPKGNPICSLHDQGAGGNG
				[		NVLKELSDPAGAIIYTSRFQLGDPTLNALEIW
		J	]			GAEYQESNALLLRSPNRDFLTHVSARERCPA
•						CFVGTTTGDRRIVLVDDRECPVRRNGQGDAP
						PTPPPTPVDLELEWVLGKMPRKEFFLQRKPP
						MLQPLALPPOLSVHQALERVLRLPAVASKRY
1			l			LTNKVDRSVGGLVAQQQCVGPLQTPLADVA
						VVALSHEELIGAATALGEQPVKSLLDPKVAA
						RLAVAEALTNLVFALVTDLRDVKCSGNWM
						WAAKLPGEGAALADACEAMVAVMAALGVA
						VDGGKDSLSMAARVGTETVRAPGSLVISAYA
			i			VCPDITATVTPDLKHPEGRGHLLYVALSPGQ
			1		·	HRLGGTALAQCFSQLGEHPPDLDLPENLVRA
. [					1	FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM
						AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV
			[			LEVQEPDLAQVLKRYRDAGLHCLELGHTGE
					·	AGPHAMVRVSVNGAVVLEEPVGELRALWEE
			<b>,</b>			TSFQLDRLQAEPRCVAEEERGLRERMGPSYC
ľ	Ì		<b>'</b>			LPPTFPKASVPREPGGPSPRVAILREEGSNGDR
			<b>i</b>			EMADAFHLAGFEVWDVTMQDLCSGAIGLDT
			·	ſ	. [	FRGVAFVGGFSYADVLGSAKGWAAAVTFHP
					]	RAGAELRRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNFDA AFMGPDSODA PROLITI PUNT
		:				WVGGDPNEDAAEMGPDSQPARPGLLLRHNL SGRYESRWASVRVGPGPALMLRGMEGAVLP
ŀ	1	'	(		ł	VWSAHGEGYVAFSSPELQAQIEARGLAPLHW
ŀ						ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR
ĺ	ſ			•	i	HLAVMPHPERAVRPWQWAWRPPPFDTLTTS
						PWLQLFINARNWTLEGSC
1345	2695	· A	10396	65	642	GVRGFWAGTMASRAGPRAAGTDGSDFOHRE
Ì	ł				- :	RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV
j				=	İ	AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI
. /	i			.	ł	LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI
	. [			ŀ		YGSMEMFPAAQQLYRHGKAYRFLFGFSAVSI
.	ľ			ļ	- 1	MYLVLVLAVQVHAWQLYYSKKLLDSWFTST
				1	1	QEKKHK
1346	2696	Α	10398	1	718	DDFVRCGPQSAAMGASARLLRAVIMGAPGS
				Į.	1	GKGTVSSRITTHFELKHLSSGDLLRDNMLRGT
			1	ļ		EIGVLAKAFIDQGKLIPDDVMTRLALHELKNL
1	1			ł	ì	TQYSWLLDGFPRTLPQAEALDRAYQIDTVINL
	į			ł		NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK
	l				•	TVGIDDLTGEPLIQREDDKPETVIKRLKAYED
Į	J			l l	·	<b>QTKPVLEYYQKKGVLETFSGTETNKIWPYVY</b>
				i		AFLQTKVPQRSQKASVTP
1347	2697	A	10402	153	1969	KHRQENNALDMAPEIHMTGPMCLIENTNGEL
,	,	,	40702 1	123 1	ו עסענ	ARKUENNALDMAPEIHMIGPMCIJENTNGW

SEQ ID	SEQ ID	Met	SEQ	Predicted	David and	[ A - ! ! ] (A - 4)
NO: of	NO: of	hod	ID NO:	beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1 1100	in NO.	nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide		1	USSN		location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-			location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
neuce	[	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		ŀ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	ł	ł	ł	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide		/=possible nucleotide deletion, \=possible
			L	sequence		nucleotide insertion
						LMNKLAGKNKGFSLGSTVKSHTKGIWMWCV
<b>.</b>	l	İ	J .			PHPKKPEHTLVLLDTEGLGDVKKGDNONDS
						WIFTLAVLLSSTLVYNSMGTINQQAMDQLYY
	i					VTELTHRIRSKSSPDENENEDSADFVSFFPDFV
1		Ì.	í			WTLRDFSLDLEADGOPLTPDEYLEYSLKLTO
			}			GTSOKDKNFNLPRLCIRKFFPKKKCFVFDLPI
i i			ľ	•		HRRKLAQLEKLQDEELDPEFVQQVADFCSYI
1			l i			FSNSKTKTLSGGIKVNGPRLESLVLTYINAISR
l l			ł i			GDLPCMENAVLALAQIENSAAVQKAIAHYD
						QMGQKVQLPAETLQELLDLHRVSEREATEV
Į l						YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK
1				-		ONOEASSDRCSALLOVIFSPLEEEVKAGIYSK
1						
	·					PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL
1						QTYLKSKESVTDAILQTDQILTEKEKEIEVEC
i						VKAESAQASAKMVEEMQIKYQQMMEEKEKS
1	ŀ					YQEHVKQLTEKMERERAQLLEEQEKTLTSKL
ł						QEQARVLKERCQGESTQLQNEIQKLQKTLKK
1348	2698		10101			KTKRYMSHKLKI
1348	2098	Α	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG
]	_					VAGGAPRRRTPVTMWRLLARASAPLLRVPLS
1					1	DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL
						RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT
1 1				1		EGNFAILALGGGYLHWGHFEMMRLTINRSM
1				ļ		DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK
1 1						GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG
						FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE
1 1				ļ		RERNNONPWTFERIATANMLGIRKVLSPYDL
						THKGKYWGKFYMPKRV
1349	2699	Α	10409	59	1184	LRRNCSALGGLFQTIISDMKGSYPVWEDFINK
				J		AGKLQSQLRTTVVAAAAFLDAFQKVADMAT
j						NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS
						ALIDCLINPLQEQMEEWKKVANQLDKDHAK
1 1						EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ
[ [		ĺ		ď		PQLDSALQDVNDKYLLLEETEKQAVRKALIE
						ERGRECTEISMLRPVIEEEISMLGEITHLQTISE
			'			DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS
j · j				j		YOTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS
; l						GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH
		l		ļ		DSGFISQDAFQSKSPSPMPPEAFNORRKEKRE
[ [				i i		PDPNGGGPTTASGPPAAAEEAORPRSM
1350	2700	A	10410	511	958	AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST
				<b>-</b>		RGHSSLLPPSODFVAGLSVILRGTVDDRLNW
}		-				AFNLYDLNKDGCITKEEMLDIMKSIYDMMG
		1			ì	KYTYPALREEAPREHVESFFQKMDRNKDGV
			;	ļ		
						VTIEEFIESCQKDENIMRSMQLFDNVI

## WHAT IS CLAIMED IS:

- 1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, and complementary sequences thereof.
- 2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
- 3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
- 4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
- 5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
- 6. A vector comprising the polynucleotide of claim 1.
- 7. An expression vector comprising the polynucleotide of claim 1.
- 8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
- 9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
- 10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
  - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
  - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO:1-1350.
- 11. A composition comprising the polypeptide of claim 10 and a carrier.
- 12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

- a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
- b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
- 14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
- b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
- c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
- 15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
- 16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
- b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
- 17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:
- a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 19. A method of producing the polypeptide of claim 10, comprising,
- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-1350, under conditions sufficient to express the polypeptide in said cell; and
  - b) isolating the polypeptide from the cell culture or cells of step (a).
- 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1351-2700, the mature protein portion thereof, or the active domain thereof.
- 21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
- 22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-1350.
- 23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
- 24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
- 25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
- 26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

Pages 340 to 1963 of this application contain amino acid sequence listings. They can be obtained at the address given below.

Les pages 340 to 1963 de cette demande contiennent des listages des séquences d'acides aminés. Elles peuvent être obtenues à l'adresse indiquée ci-dessous.

World Intellectual Property Organization 34, chemin des Colombettes CH-1211 Genève 20

## This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

BLACK BORDERS

IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

FADED TEXT OR DRAWING

BLURRED OR ILLEGIBLE TEXT OR DRAWING

SKEWED/SLANTED IMAGES

COLOR OR BLACK AND WHITE PHOTOGRAPHS

GRAY SCALE DOCUMENTS

LINES OR MARKS ON ORIGINAL DOCUMENT

REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

## IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.